

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 06:18:03 ; Search time 75 Seconds  
(without alignments)  
2158.663 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910

Sequence: 1 MITALDLVHLTVAVPLVYA.....DILSTAYGPTSGHGFITCHS 573

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A: Genesep29Jan04:\*

1: Genesep1980s:\*

2: Genesep1990s:\*

3: Genesep2000s:\*

4: Genesep2001s:\*

5: Genesep2002s:\*

6: Genesep2003as:\*

7: Genesep2003bs:\*

8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2910	100.0	573	3 AAB26934	Corn auxi
2	1665	57.2	595	3 AAY44265	Aay44265 Rice EIR1
3	1647.5	56.6	622	5 ABB91576	Abb91576 Herbicide
4	1643	56.5	605	3 AAB26942	Aab26942 Soybean a
5	1641.5	56.4	622	3 AAB26950	Aab26950 Auxin tra
6	1562.5	53.7	632	3 AAB26946	Aab26946 Wheat aux
7	1560.5	53.6	640	3 AAG35877	Aag35877 Arabidops
8	1560.5	53.6	640	5 ABB91542	Abb91542 Herbicide
9	1542.5	53.0	616	3 AAG31983	Aag31983 Arabidops
10	1542.5	53.0	616	5 ABB91685	Abb91685 Herbicide
11	1541	53.0	647	3 AAY44270	Aay44270 A. thalia
12	1535	52.7	647	3 AAY53129	Aay53129 Gravitrop
13	1535	52.7	647	3 AAY44264	Aay44264 A. thalia
14	1535	52.7	647	3 AAY44271	Aay44271 A. thalia
15	1535	52.7	647	3 AAB26949	Aab26949 Auxin tra
16	1533	52.7	647	3 AAY44272	Aay44272 A. thalia
17	1528	52.5	589	3 AAB26937	Aab26937 Rice auxi
18	1527	52.5	637	3 AAB26944	Aab26944 Soybean a
19	1526	52.4	619	3 AAG36342	Aag36342 Arabidops
20	1526	52.4	619	5 ABB91105	Abb91105 Herbicide
21	1491.5	51.3	620	3 AAG35878	Aag35878 Arabidops
22	1468.5	50.5	596	3 AAG31984	Aag31984 Arabidops
23	1453	49.9	599	3 AAG36343	Aag36343 Arabidops
24	1408.5	48.4	574	3 AAG40525	Aag40525 Arabidops
25	1406.5	48.3	660	5 ABB93849	Abb93849 Herbicide

#### ALIGNMENTS

##### RESULT 1

AAB26934

ID AAB26934 standard; protein; 573 AA.

XX AAB26934;

AC AAB26934;

XX 02-FEB-2001 (first entry)

XX Corn auxin transport protein clone p0119.cmtnl24r.

XX Auxin transport protein; corn; root development; gene mapping;

XX plant breeding; herbicide.

XX Zea mays.

XX W0200608389-A2.

XX 16-NOV-2000.

XX 03-MAY-2000; 2000WO-US012061.

XX 07-MAY-1999; 99US-0133040P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX (PION-) PIONEER HI-BRED INT INC.

XX Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;

XX WPI; 2000-687647/67.

XX N-PSDB; AAA94721.

XX New nucleic acid sequences encoding new auxin transport proteins, useful for modulating root growth of plants and to screen for herbicides.

XX Claim 10; Fig 1; 94pp; English.

XX Auxins are plant hormones that influence plant behaviour and development e.g. vascular tissue differentiation, apical development, tropic responses and organ (e.g. flower, leaf) development. The present invention relates to auxin transport proteins and coding sequences. The present sequence is one such auxin transport protein. The DNA sequence encoding this protein may be used to modulate root development, e.g. to produce a more robust root system, alter root angle or redirect root growth; and may be useful for gene mapping (e.g. for plant breeding) and to identify loss of function mutants. This protein may be useful for raising specific antibodies, for the detection of auxin transport proteins and to design and/or identify specific inhibitors of auxin transport proteins, potentially useful as herbicides

```

XX Sequence 573 AA;
SQ
Query Match
Best Local Similarity 100.0%; Score 2910; DB 3; Length 573;
Matches 573; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MITALDLVHLTAIVVPLVAVMTLAYGVSRVWRRIPTDQCSCGINRFVALFAVPLLSFHFIS 60
DB 1 MITALDLVHLTAIVVPLVAVMTLAYGVSRVWRRIPTDQCSCGINRFVALFAVPLLSFHFIS 60

QY 61 TNDPFAMNRLFLAATLOKVAIVLALALAGSLSPRALGLDWSITLFSITLPTLVMG 120
DB 61 TNDPFAMNRLFLAATLOKVAIVLALALAGSLSPRALGLDWSITLFSITLPTLVMG 120

QY 121 IPLLGRMTGASAGTLMVQVVLQCIITWYTLMLFLFYRAARALVLDQPDGAAASIVSF 180
DB 121 IPLLGRMTGASAGTLMVQVVLQCIITWYTLMLFLFYRAARALVLDQPDGAAASIVSF 180

QY 181 RVDSDVSLARGDVELEAPDGVAGAGAVSRGSDAGRVVTVRKSTSSRSEAAASHSHS 240
DB 181 RVDSDVSLARGDVELEAPDGVAGAGAVSRGSDAGRVVTVRKSTSSRSEAAASHSHS 240

QY 241 QTMQPRVNLSGVEIYLSQSSRNPTPRGSSFNHADFNIVGAAGKGGGAAGDEEKGACG 300
DB 241 QTMQPRVNLSGVEIYLSQSSRNPTPRGSSFNHADFNIVGAAGKGGGAAGDEEKGACG 300

QY 301 GGGGSHSPQQAQAVAPAKRDLHMLVWSSASPVSEAAVHVFGAGADHADVLAAGAQA 360
DB 301 GGGGSHSPQQAQAVAPAKRDLHMLVWSSASPVSEAAVHVFGAGADHADVLAAGAQA 360

QY 361 YDEYGRDDYSRTKNGSGGADKGGPTLSKLGSNSTAQLYPKDDGEGRAAAMPASVMT 420
DB 361 YDEYGRDDYSRTKNGSGGADKGGPTLSKLGSNSTAQLYPKDDGEGRAAAMPASVMT 420

QY 421 RLILIMWRKLRNPNTYSSILGVVWSLVSYRNGIEMPAIIARSISILSDAGLGMAMFSL 480
DB 421 RLILIMWRKLRNPNTYSSILGVVWSLVSYRNGIEMPAIIARSISILSDAGLGMAMFSL 480

QY 481 GLFWALQRIIACGNKLAAMGVRFVAGPAVMAAASIAVLGVLHLLTAIVQALPOQI 540
DB 481 GLFWALQRIIACGNKLAAMGVRFVAGPAVMAAASIAVLGVLHLLTAIVQALPOQI 540

QY 541 VPFVFAKEYGVHPDILTAYGPTTSHGFITCHS 573
DB 541 VPFVFAKEYGVHPDILTAYGPTTSHGFITCHS 573

RESULT 2
RAY44265
ID AAY44265 standard; protein; 595 AA.
AC AAY44265;
XX
DT 28-FEB-2000 (first entry)
DE Rice EIR1 Homologue (REH1) protein.
KW Root-specific auxin transport protein; EIR1; Ethylene Insensitive Root;
KW gravitropism; auxin-based herbicide; plant root; transgenic plant;
KW auxin homeostasis; Rice EIR1 Homologue; REH1; transmembrane domain.
XX
OS Oryza sativa.
XX
FH Key
FH Peptide
FT Location/Qualifiers
FT 1..60
FT /label= Signal_peptide
FT 10..28
FT /label= Transmembrane_domain
FT 41..60
FT /label= Transmembrane_domain
FT 71..92
FT /label= Transmembrane_domain

```

```

FT Domain 100..121
FT /label= Transmembrane_domain
FT Domain 132..152
FT /label= Transmembrane_domain
FT Domain 451..468
FT /label= Transmembrane_domain
FT Domain 478..501
FT /label= Transmembrane_domain
FT Domain 516..537
FT /label= Transmembrane_domain
FT Domain 543..559
FT /label= Transmembrane_domain
FT Domain 570..590
FT /label= Transmembrane_domain
XX
PN WO9963092-A1.
XX
XX 09-DEC-1999.
XX
XX 03-JUN-1999; 99WO-US012277.
XX
XX 03-JUN-1998; 98US-0087789P.
XX
XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
XX Luschnig C, Gaxiola RA, Grisafi P, Fink GR;
XX
XX WPI; 2000-086979/07.
XX
XX N-PSDB; AAZ29280.
XX
XX DNA encoding a root-specific auxin transport protein, used to develop
XX transgenic plants with increased resistance to herbicides.
XX
XX Claim 18; Fig 9; 55pp; English.
XX
XX The present sequence is a REH1 protein, a rice homologue of EIR1 which
XX functions as a root-specific auxin transport (efflux) protein involved in
XX gravitropism. The sequence is obtained from a rice EST derived from root-
XX specific cDNA. REH1 consists of potential N-glycosylation sites and a N-
XX terminal signal peptide. The transmembrane domains are located in the
XX highly conserved portions of the protein N- and C-termini. The sequence
XX is used for producing genetically engineered plants with greater
XX resistance to auxin-based herbicides and auxin transport inhibitors in
XX combination with a second herbicide. It can also be used to enhance
XX transport of auxin in plant roots, produce transgenic plants which
XX exhibit altered auxin homeostasis and mutant plants in which the roots
XX are agravitropic and have reduced sensitivity to ethylene
XX
XX Sequence 595 AA;
SQ
Query Match 57.2%; Score 1665; DB 3; Length 595;
Best Local Similarity 58.9%; Pred. No. 3e-147;
Matches 365; Conservative 54; Mismatches 111; Indels 90; Gaps 17;

QY 1 MITALDLVHLTAIVVPLVAVMTLAYGVSRVWRRIPTDQCSCGINRFVALFAVPLLSFHFIS 60
DB 1 MITADDFYHVMITAMVPLVAVMTLAYGVSRVWRRIPTDQCSCGINRFVALFAVPLLSFHFIS 60

QY 61 TNDPFAMNRLFLAATLOKVAIVLALALAGSLSPRALGLDWSITLFSITLPTLVNT 116
DB 61 TNNPYTNLRFIAATLOKVAIVLALALAGSLSPRALGLDWSITLFSITLPTLVNT 113

QY 117 LVNMGIPLLRGMTGASAGTLMVQVVLQCIITWYTLMLFLFYRAARALVLDQPDGAAAS 176
DB 114 LVNMGIPLLRGMTGAFS-GSLMVQIVVLQCIITWYTLMLFMFYRGARMLITEQFPD-TAAN 171

QY 177 IVSFRVDSVLSL--ARGDVELEAPDGVAGAGAVSRGSDAGRVVTVRKSTSSRSEAA 234
DB 172 IASIVVDVDSVLSLGRDRAITETEVK-----EDGRHVTVRKSNASRSIY 218

QY 235 CSHSHS-QTMQPRVNLSGVEIYLSQSSRNPTPRGSSFNHADFNIVGAAGKGGGAAGD 293
DB 219 SRSMGFSSTTPRPSNLTNAEYLSQSSRNPTPRGSSFNHADFNIVGAAGKGGGAAGD 278

```

QY 294 EEKGAC-----GGGGGH--SPOQAVAVP-----AKR 319  
 Db 279 VRTGATPRPSNYEDDASKPKYPLPASNAAPMAGHYPAFNPVSSAPKGAKAATNGQAKG 338  
 QY 320 KDLHMLWSSSSASPVSERAAVHVFAGGADHADVLA-----KGAQAYDEY-GRDDYSS 371  
 Db 339 EDLHMFVSSSSASPVSD-----VFGGAPDYNDAAAIVKSPRKDGAKOREDYVEDDFSF 393  
 QY 372 RTKNGSGGADKGGPTLSKLSNSTAQLYPKDDEGGRAAAAMVAPPPASVMTLILIMVWKL 431  
 Db 394 GNR-----GVMDRD-----AEAGDEKAAAAGADPSKAWAAPTAMPTSVMTLILIMVWKL 446  
 QY 432 IRNPNTYSSLLGVWVSVRWGCIEMPAIIARSISILSDAGLGMAFSLGLFMALQPII 491  
 Db 447 IRNPNTYSSLLGVWVSVRWGCIEMPAIIARSISILSDAGLGMAFSLGLFMALQPII 506  
 QY 492 ACNKLAAIAMGVRFVAGPAVMAAASIAVGLRGLLHIAIQAALPQGIIVPFVFAKEYGV 551  
 Db 507 ACNKKVATYAWARFLAGPAVMAAASFAVGLRGLLHIAIQAALPQGIIVPFVFAKEYSV 566  
 QY 552 HPDILSTA--YG-----PIT 564  
 Db 567 HPSILSTAVIFGMLIALPIT 586

## RESULT 3

ID ABB91576  
 XX ABB91576 standard; protein; 622 AA.  
 AC ABB91576;  
 XX  
 DT 31-MAY-2002 (first entry)  
 XX  
 DE Herbicidally active polypeptide SEQ ID NO 787.  
 XX  
 KW Herbicidal; plant; agriculture; herbicide.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FN WO200210210-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 28-AUG-2001; 2001WO-EP009892.  
 XX  
 PR 28-AUG-2001; 2001WO-EP009892.  
 XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Tietjen K, Weidler M;  
 XX  
 DR WPI; 2002-269010/31.  
 XX

XX Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.  
 XX  
 PS Claim 5; SEQ ID NO 787; 261pp + Sequence Listing; English.  
 XX  
 CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for  
 CC identifying modulators. The identified modulators are useful as  
 CC herbicides  
 XX  
 SQ Sequence 622 AA;

Query Match 56.6%; Score 1647.5; DB 5; Length 622;  
 Best Local Similarity 56.2%; Pred No. 14e-145;  
 Matches 365; Conservative 56; Mismatches 104; Indels 125; Gaps 18;  
 QY 1 MITALDLYHVLTAIVPLIYVAMTLAGYSVRWKRIPTDQCSGINRFVALFVAPLLSFHIS 60  
 Db 1 MITAADFHYHMTAMVPLIYVAMTLAGYSVRWKRIPTDQCSGINRFVALFVAPLLSFHIA 60  
 QY 61 TNDPPFAMNLFRLAADTLQKVAVALIALALASRLGSLSPRALGLDWSITLPSLSTLPNTLVMG 120  
 Db 61 ANNPYAMNLFRLAADSLQKIVLISLLFLWCK--LS--RNGSLDWTITLPSLSTLPNTLVMG 117  
 QY 121 IPLLGRMYGASAGTLMQVIVVLCIIWYTLMLFLFEYRAAPALVLDOPFGAASISVF 180  
 Db 118 IPLLKMGYG-NRSGDLMQVIVVLCIIWYTLMLFLFEYRGAKLLISEQFPD-TAGSIYSI 175  
 QY 181 RVSDVVSIV-ARGDVELEAEPDGVAGAGAVSRGSDAGRVVTRKSTSSRSEACSHSH 239  
 Db 176 HVSDSDIMSLDGRQPLETAEIK-----EDGKLHVTVERSNASHSDIYSRSQ 222  
 QY 240 SQTMQPRVSNLGSVEIYISLOSSRNPTRGSSFNHADPFNIV----- 280  
 Db 223 GLSATPRPSNLNAEIIYISLOSSRNPTRGSSFNHADPFNIV----- 282  
 QY 281 -----GAAAK--GGGGAAG-----DEEKGACGGGGGGHSPOP----- 310  
 Db 283 KGTPRPSNVEEDGGPAKETAAGTAAGRFFHYQSGSGGGGGAHYPAFNPQMFSPNTGG 342  
 QY 311 -----QAVVPAKR-----KDLHMLWSSSSASPVSERAAVHVFAGGADHADVLA 357  
 Db 343 GGGTAAKGNAPVVGGRQDNGRDLHMFVWSSSSASPVSD-----VFGGGGNGHH---ADY 394  
 QY 358 AQAYDEYGRD-----DYSSRTKXSGGADKGGPTLSKLG-----SNSTAQLYP 400  
 Db 395 STATNDHQDKVKISVPQNSNDNQVVEREEFSGNKKDDSKVLATDGGNNISNKTQ--- 451  
 QY 401 KDCGEGRAAAVAMPASPASVMTLILIMVWKLIRNPNTYSSLLGVWVSVRWGEMPAI 460  
 Db 452 -----AKYMPPTSVMTLILIMVWKLIRNPNTYSSLLGVWVSVRWGEMPAI 502  
 QY 461 IARISILSDAGLGMAFSLGLFMALQPIIACGNKLAIAIAGVRFVAGPAVMAAASIAV 520  
 Db 503 IAKSISILSDAGLGMAFSLGLFMALQPIIACGNRRRAAFAAANRFVVGPAVMLVASVAV 562  
 QY 521 GLRGVLLHIAIQAALPQGIIVPFVFAKEYGVHPDILSTA--YG-----PIT 564  
 Db 563 GLRGVLLHIAIQAALPQGIIVPFVFAKEYGVHPDILSTAVIFGMLIALPIT 613

## RESULT 4

AAB26942  
 ID AAB26942 standard; protein; 605 AA.  
 XX  
 AC AAB26942;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX  
 DE Soybean auxin transport protein clone sfl1.pk131.g9.  
 XX  
 KW Auxin transport protein; soybean; root development; gene mapping;  
 KW plant breeding; herbicide.  
 XX  
 OS Glycine max.  
 XX  
 PN WO2000068389-A2.  
 XX  
 PD 16-NOV-2000.  
 XX  
 PF 03-MAY-2000; 2000WO-US012061.  
 XX  
 PR 07-MAY-1999; 99US-0133040P.  
 XX  
 PA (DUPO ) DU PONT DE NEMOURS & CO E I.

PA (PION-) PIONEER HI-BRED INT INC.  
 XX Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
 XX WPI; 2000-687647/67.  
 DR N-PSDB; AAA94729.  
 XX New nucleic acid sequences encoding new auxin transport proteins, useful  
 PT for modulating root growth of plants and to screen for herbicides.  
 XX Claim 10; Fig 1; 94pp; English.  
 XX Auxins are plant hormones that influence plant behaviour and development  
 CC e.g. vascular tissue differentiation, apical development, tropic  
 CC responses and organ (e.g. flower, leaf) development. The present  
 CC invention relates to auxin transport proteins and coding sequences. The  
 CC present sequence is one such auxin transport protein. The DNA sequence  
 CC encoding this protein may be used to modulate root development, e.g. to  
 CC produce a more robust root system, alter root angle or redirect root  
 CC growth; and may be useful for gene mapping (e.g. for plant breeding) and  
 CC to identify loss of function mutants. This protein may be useful for  
 CC raising specific antibodies, for the detection of auxin transport  
 CC proteins and to design and/or identify specific inhibitors of auxin  
 CC transport proteins, potentially useful as herbicides  
 XX Sequence 605 AA;  
 XX  
 XX Query Match 56.5%; Score 1643; DB 3; Length 605;  
 XX Best Local Similarity 57.3%; Pred. No. 3.6e-145;  
 XX Matches 361; Conservative 58; Mismatches 111; Indels 100; Gaps 17;  
 QY 1 MITALDLVHVLTAUVPLVYAMTLAGSVRWRIPTDQCSCGINRFVALFAVPLLSFHFI 60  
 DB 1 MITLDFHVTAMVPLVYAMTLAGSVKWKIFSPDQCSCGINRFVALFAVPLLSFHFI 60  
 QY 61 TNDPFAMNLFRAADTLQKVAVLALALASRGLSSPRALGLDWSITLPSLSTPLNT 116  
 DB 61 SNPYENMLFLAADTLQKIILVLLAVMSNITKRG-----LEWATLPSLSTPLNT 113  
 QY 117 LVMGPIILRMGMGASSAGTLMVQVVLQCIITWTMLFPEYRAARALVLDDPPDGAAS 176  
 DB 114 LVMGPIILRMGMGDFPS-GSLMVQVVLQCIITWTMLFLEFGRARWLISEQPD-TAAS 171  
 QY 177 IVSFVSDVSVL-ARGDVELEAPDPGAGAVSVSRGGDAGRVVTRKSTSSRSEAC 235  
 DB 172 IVSIVSDVSVMLDGRQPLETEAIK-----EDGKLHVTVRKSNASRSDIFS 218  
 QY 236 SHSHS-QTMQPRVSNLGVETYSLQSSRNPTPRGSSFNHADFFNIYGAAKGGGGAAGD- 293  
 DB 219 RSQGLSTTPRPNLTNAEIYSLQSSRNPTPRGSSFNHDTFYSMAAGRSNFCASDV 278  
 QY 294 -----BEKGA-----CGGGGGHSPQO-----AVAPAK 318  
 DB 279 YGLSASRGFTPRPSNYDEGGKPFHYHAAGTGHYPAPNPGMFSPSNGSKSYAANNAK 338  
 QY 319 R-----KDLHMLVSSSSASPVSEARAAPHVFGAGGADHADVLAKGAQAYDEYG 365  
 DB 339 RPNQAQLKPDGHRDLHMFVSSSSASPVSDVFGAHEYG-CGHQKEVKLNVPSPKVEYN 397  
 QY 366 ----RDDYSTKXGSGGADGGPTLSKLSGNSNTAQLYPKDDGEGRAAAVAMPASVMT 421  
 DB 398 HRDTQEDYLEKDEFSFGNREM-----DRENNQLEGEKVGDKPK-TWPPASVMT 446  
 QY 422 LILVWVKLRNPNNTYSSLLGVVSVSVRWKGTWEMPAIARSISILSDAGLGMAMFSLG 481  
 DB 447 LILVWVKLRNPNNTYSSLLGLTWSLVSPKXNVEMPAIARSISILSDAGLGMAMFSLG 506  
 QY 482 LFMALQPRIIACGNKLAANGVRFVAGPAMAAASIAVGLRGVLLHIAIVQALPQGV 541  
 DB 507 LFMALQPRVIAACGNSTAFANVRLTGPAMAAASIAVGLRGVLLHIAIVQALPQGV 566  
 QY 542 PVPFAKEYGVHDPDILSTA--YG-----PIT 564

DB 567 PTFVAKYNNVHPDILSTAVIFGMLIALPIT 596  
 RESULT 5  
 AAB26950  
 ID AAB26950 standard; protein; 622 AA.  
 XX  
 AC AAB26950;  
 XX  
 DT 02-FEB-2001 (first entry)  
 XX  
 DE Auxin transport protein AtPIN1.  
 XX  
 KW Auxin transport protein; corn; root development; gene mapping;  
 XX plant breeding; herbicide.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 PN WO200068389-A2.  
 XX  
 PD 16-NOV-2000.  
 XX  
 PF 03-MAY-2000; 2000WO-US012061.  
 XX  
 PR 07-MAY-1999; 99US-0133040P.  
 XX  
 PA (DUPO) DU PONT DE NEMOURS & CO E I.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 XX  
 PI Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
 XX  
 XX WPI; 2000-687647/67.  
 DR  
 XX  
 PT New nucleic acid sequences encoding new auxin transport proteins, useful  
 PT for modulating root growth of plants and to screen for herbicides.  
 XX  
 PS Disclosure; Fig 1; 94pp; English.  
 CC  
 CC Auxins are plant hormones that influence plant behaviour and development  
 CC e.g. vascular tissue differentiation, apical development, tropic  
 CC responses and organ (e.g. flower, leaf) development. The present  
 CC invention relates to auxin transport proteins and coding sequences. The  
 CC present sequence is one such auxin transport protein. The DNA sequence  
 CC encoding this protein may be used to modulate root development, e.g. to  
 CC produce a more robust root system, alter root angle or redirect root  
 CC growth; and may be useful for gene mapping (e.g. for plant breeding) and  
 CC to identify loss of function mutants. This protein may be useful for  
 CC raising specific antibodies, for the detection of auxin transport  
 CC proteins and to design and/or identify specific inhibitors of auxin  
 CC transport proteins, potentially useful as herbicides  
 XX Sequence 622 AA;  
 XX  
 XX Query Match 56.4%; Score 1641.5; DB 3; Length 622;  
 XX Best Local Similarity 56.1%; Pred. No. 5.2e-145;  
 XX Matches 365; Conservative 56; Mismatches 105; Indels 125; Gaps 18;  
 QY 1 MITALDLVHVLTAUVPLVYAMTLAGSVRWRIPTDQCSCGINRFVALFAVPLLSFHFI 60  
 DB 1 MITAADFVHTAMVPLVYAMTLAGSVKWKIFTPDQCSCGINRFVALFAVPLLSFHFI 60  
 QY 61 TNDPFAMNLFRAADTLQKVAVLALALASRGLSSPRALGLDWSITLPSLSTPLNTLYMG 120  
 DB 61 ANNPYAMNLFRAADSLQKVLVLSLFLWCK-LS-RNGSLDWTITLPSLSTPLNTLYMG 117  
 QY 121 IPLLGRMGYASSAGTLMVQVVLQCIITWTMLFPEYRAARALVLDDPPDGAASIVSF 180  
 DB 118 IPLLGRMGY-NFGSDLMVQVVLQCIITWTMLFPEYRAAKLLISEQPD-TAGSIVSI 175  
 QY 181 RVSDVSVL-ARGDVELEAPDPGAGAVSVSRGGDAGRVVTRKSTSSRSEACSHH 239  
 DB 176 HVSDIMSLDGRQPLETEAIK-----EDGKLHVTVRKSNASRSDIYSRSRQ 222



XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130445P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135623P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 22-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.

PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 18-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144864P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 25-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.

PR 22-SEP-1999; 99US-01551399.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 25-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 53.6%; Score 1560.5; DB 3; Length 640;  
Best Local Similarity 52.9%; Pred. No. 2.2e-137;  
Matches 353; Conservative 73; Mismatches 102; Indels 139; Gaps 20;

QY 1 MITALDLHYLTAVPLVYAMTILAYGVNRWRIPTDQCSGINRFVALFAVPLLSFHFIS 60  
DB 1 MISHWDLTYLTAVPLVYAMTILAYGVNRWRIPTDQCSGINRFVALFAVPLLSFHFIS 60  
QY 61 TNDPFAMNLFRAADTLQKAVLALALASRLGLSDWLTSLTSLTNTLVWG 120  
DB 61 TNNPFAMNLFRAADTLQKAVLALALASRLGLSDWLTSLTSLTNTLVWG 117  
QY 121 IPLLGRWYGASSAGTLMVQVVLQCTIIVYTLMLFLFEYRAAEALVLDQPPGAAASIVSF 180  
DB 118 IPLLIMYGEYS-GSLMVQIVLQCTIIVYTLMLFLFEYRAAEALVLDQPPGAAASIVSF 175  
QY 181 RVDSVDVVLARGD-VELEAEPDGVAGAGAVSRGGDAGRVVTVRKTSRSEACSHS 239  
DB 176 KVESDVVSLDGHDFLETDAEI-----GDDGKLHVTVRKSNASR-RSFC--- 217  
QY 240 SQTMQVRVNLQSVETYSQSRNPTPGSSFNHADFNIVG-----A 282  
DB 218 GPNWTPRPNLTGAELYSJST-----TPRGSNNHSDFTNMWGFPGRLSNFGPADMSVQ 273  
QY 283 AAKG-----GGGAAGDEKGAACGGCGGCHSPQ----- 310  
DB 274 SSRGTPRPSNFENECAMASSPRFGYPPGGGAGSYPAENPFRFSSTTTSTANKSVNKNPKD 333  
QY 311 ----QAVAVPA-----KXKDLHMLVSSSPYSERAAVHVFCAGGADR-----ADVL 354  
DB 334 VNTNQTTTLTGKSNSHDAKELHFWSSNGSPVSDRAGLVFG-GAPDNDQGRSDQG 392  
QY 355 AK-----GAQAYDEYGRDDYSSTRTKNGSGADKGGFTL 387

DB 393 AKBIRMLVPDQSHNGETKAVAHPASGDGQEQFQFAGKEEAERPKDAENG-----L 445  
QY 388 SKLGSNSTAQLPKDQDEGARAAVA---MPPASVNTRLILIMVWKLIRNNTYSSLIGV 444  
DB 446 NKLAPNSTAALQSK-TGLGGAASQKMPASVNTRLILIMVWKLIRNNTYSSLIGL 504  
QY 445 VWSLVSRWGIENPAILIARSISILSDAGLGMAWFSGLFLMALQPRIIACGNKLAIAIMGV 504  
DB 505 IVALVAFRWHVAMPKTIQOSISILSDAGLGMAWFSGLFLMALQPKLIACGNSVATFAMAV 564  
QY 505 RFVAGPAVMAAASIAVGLRGVLLHAIQAALPQGIIVPFAKEVGVHPDILSTA--YG- 561  
DB 565 RFLTGPAVMAVAIAIGLRGDLRLVAIVQAALPQGIIVPFAKEVGVHPDILSTA--YG- 624  
QY 562 ----PIT 564  
DB 625 LTALPIT 631

RESULT 8  
ABB91542  
ID ABB91542 standard; protein; 640 AA.

XX ABB91542;  
XX 31-MAY-2002 (first entry)  
XX Herbicidally active polypeptide SEQ ID NO 753.  
XX Herbicidal; plant; agriculture; herbicide.  
XX Arabidopsis thaliana.  
XX WO200210210-A2.  
XX 07-FEB-2002.  
XX 28-AUG-2001; 2001WO-EP009892.  
XX 28-AUG-2001; 2001WO-EP009892.  
XX (FARB ) BAYER AG.  
XX Tietjen K, Weidler M;  
XX WPI; 2002-269010/31.  
XX

Identifying plant target proteins for herbicidally active compounds,  
comprising aligning and comparing nucleic acid or amino acid sequences  
from plant with nucleic acid or amino acid sequences from non-plant  
organisms.  
Claim 5; SEQ ID NO 753; 261pp + Sequence Listing; English.

The invention relates to identifying target proteins (ABB90790-ABB94016)  
for herbicidally active compounds, comprising aligning and comparing  
nucleic acid or amino acid sequences from plant with nucleic acid or  
amino acid sequences from non-plant organisms using suitable search  
parameters, where plant sequences having an E-value greater by a factor  
of 3 than the E-value of most similar non-plant sequences are selected.  
The polypeptides or nucleic acids encoding them are useful for  
identifying modulators. The identified modulators are useful as  
herbicides

Sequence 640 AA;  
Query Match 53.6%; Score 1560.5; DB 5; Length 640;  
Best Local Similarity 52.9%; Pred. No. 2.2e-137;  
Matches 353; Conservative 73; Mismatches 102; Indels 139; Gaps 20;  
QY 1 MITALDLHYLTAVPLVYAMTILAYGVNRWRIPTDQCSGINRFVALFAVPLLSFHFIS 60

Db 1 MTSWHDLYTLVTAIPLVYVAMILAYGVSRVWKIFSPDQCSGINRFAVFAVPELLSFHIS 60  
QY 61 TNDPFAMLRFLAADTLQKAVVALLALASRGLSPRALGLDWSITLPSLTPLNTLYMG 120  
Db 61 TNNPFAMLRFLAADTLQKIMLSLVLWA---NFRSGSLEWSTIFSLSTPLNTLYMG 117  
QY 121 IPLLKMYGASSAGTLMVGVVWLQCIITWTLMLFLFEYRAARALVDQPPGAAASIVSF 180  
Db 118 IPELLANYGEYS-GSLMVOIVLQCIITWTLMLFLFEYRAARALVDQPPGAAASIVSF 175  
QY 181 RVSDVVSIAARD-VELEAPDGVAGAGAVSRGGDAGRVRVTYKSTSSSEACSHSH 239  
Db 176 KVESDVVSLDGHDFLETAEI-----GDDGKLHVTVRKSNAASR-RSFC---- 217  
QY 240 SOTMOPRVNLSGVEIYLSQSSRNPTGRSSFNHADFFNIVG-----A 282  
Db 218 GNMTPRPNLTGAEIYLSLT-----TPRGSNFNHSDFYNNMGFPGRSLNFGPADMYSVQ 273  
QY 283 AAKG-----GGGAAGDEKAGCGGGGGHSPQ----- 310  
Db 274 SSRGPTPRPSNFENCAMASSPRFGYPGGAGSYPAHPNPFSSTTTSTANKSVNKNPKD 333  
QY 311 ----QAVVPA-----KKDLHMLVSSSASPVSERAAHVFGAGGADH----ADVL 354  
Db 334 VNTNQOTTLTGTGKSNHDAKELHMFVNSNGSPVSDRAGIIVFG-GAPDNDQGRSDQG 392  
QY 355 AK-----GAQAYDEYGRDYSSRTKXGSGGADKGGPTL 387  
Db 393 AXEIRMLVPDOSHNGETKVAHPASGDFGGEQOFSPAGKEEBAERPKDAENG-----L 445  
QY 388 SKLGSNSTACLYPKDGEGRARA---MPPASVMTLRLILMVWRKLIIRNPYSSLIQV 444  
Db 446 NKLAPNSTAALQSK-TGLGAEASQRKNMPPASVMTLRLILMVWRKLIIRNPYSSLIQV 504  
QY 445 VMSIVSYRWGIEMPAIARSISILSDAGLGMAMFSLGLFMALQPRIIACGNKLAIAIMGV 504  
Db 505 IHALVAFRWHVAMPKIIQQSISILSDAGLGMAMFSLGLFMALQPKLIACGNSVATPAVAV 564  
QY 505 RFVAGPAMVMAASIAVLGRVLHIALVQALPQGIIVPFVFAKEYGVHPDILSTA-YG- 561  
Db 565 RFLTGPAVMAVAAIATGLRGDLLEVAIVQALPQGIIVPFVFAKEYGVHPDILSTA-YG- 624  
QY 562 ----PIT 564  
Db 625 LIALPIT 631

RESULT 9  
AAG31983  
ID AAG31983 standard; protein; 616 AA.  
XX AC AAG31983;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 38502.  
XX KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; Genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 06-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139432P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.

PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0158293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144624P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144864P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145227P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	25-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149428P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150568P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	30-AUG-1999;	99US-0151080P.			
PR	31-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155133P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157863P.			
PR	07-OCT-1999;	99US-0158029P.			
Query Match 53.0%; Score 1542.5; DB 3; Length 616;					
Best Local Similarity 53.5%; Pred. No. 1e-135;					
Matches 349; Conservative 59; Mismatches 111; Indels 133; Gaps 17;					
Qy	1	MITADLXHVLTAVVPLVYVAMTILAYGVSVMWRIFFPDQCSGINRFFVALFAVPLLSFHFIS	60		
Db	1	MITWHLIYVLTAVVPLVYVAMTILAYGVSQVWVKIFSPDQCSGINRFFVAIFAVPLLSFHFIS	60		
Qy	61	TNDPFANLRFIAADTLQKVALALALASRGUSSPRALGLDWSITLFSLSLPTLTVMG	120		
Db	61	TNDPYANFRFAADTLQKIIMLVLLALWA--NLTKNGSLEWMITIFLSLSTLPTLTVMG	117		
Qy	121	IPLLRGWYGASSAGTLMVQVVLQCIITWTLMFLFEYBAARALVLDQPPDGAASIVSF	180		
Db	118	IPLLIANYG-TYAGSLMVQVVLQCIITWTLLFLFEYBAKLLINEQFPE-TGASIVSF	175		
Qy	181	RVDSDVVSLARGD-VELEAEPDVGAGAGAVSRGGDAGVRVTVRKSTSRSEACSHSH	239		
Db	176	KVESDVVSLDGHDFLETAEL-----GNDGKLHVTVRKSNAJR-----R	214		
Qy	240	SQTMQPRVNLSCVEIYSLQSSNPTRGSSFNHADFFNIVGAAKGGGAAGD-----	293		
Db	215	SLMTPRPSNLTGAETYSLS-----PPRGSNFNHSDFYSVMGFP-----GRLSNFQPADL	266		
Qy	294	-----EEKGACGGG--GGGHSFPQOAVAPAKR-----	319		
Db	267	YSVQSSRGPTFRPSNFENNAVKYGFYNNNTSSVPAAGSVAPAPNPEFSTGTGVSTKPNKI	326		
Qy	320	-----KDLHMLVWSSASPVSEAAVHVFGAGADHADVLAKGAQA	360		
Db	327	PKENQQLOEKDSKASHDAKELHMFVWSSASPVSD-----VFGGAGDNV-----ATE	375		
Qy	361	YDEYGRDDY-----SSRTKNGSGGAD-----KGGPTLSKLGNSSTAQLY	399		
Db	376	QSEQGAKEIRMVVSDQPRKSNARGGGDDIGGLDSGGERIEKATAGLNKMGNSSTAEL	435		
Qy	400	PKDDGEGRAAAVAMPASVMTLILIMVWRKLIRNPTVSSLLIGVWSLVSYRWGLEMPA	459		
Db	436	AAGDGGGNGNTHMPTTSVMTLILIMVWRKLIRNPTVSSLLIGLWALVAVRWHVAMPK	495		
Qy	460	IIARSTILSDAGLGMWFSLGLFMALQPRIACGNKLAIAIANGVRFVAGPVAJAAASIA	519		
PR	07-OCT-1999;	99US-0158029P.			

Db 496 ILQSSISILSDAGLGMFSLGLFVALQPKIIACGNSVATFAMAVRFTITGPAIMAVAGIA 555  
 QY 520 VGLRGVLLHIAIIVQALPQGIIVPFVFAKEYGVHPDILSTA--YG-----PIT 564  
 Db 556 IGLHGDLLRIAIVQALPQGIIVPFVFAKEYGVHPDILSTA--YG-----PIT 607

RESULT 10  
 ABB91685  
 ID ABB91685 standard; protein; 616 AA.  
 AC ABB91685;  
 XX  
 DT 31-MAY-2002 (first entry)  
 DE Herbicidally active polypeptide SEQ ID NO 896.  
 DE Herbicidal; plant; agriculture; herbicide.  
 KW Herbicidal; plant; agriculture; herbicide.  
 XX Arabidopsis thaliana.  
 OS Arabidopsis thaliana.  
 XX  
 PN WC200210210-A2.  
 XX  
 PD PD  
 XX 07-FEB-2002.  
 XX  
 XX 28-AUG-2001; 2001WO-EP009892.  
 PF  
 XX 28-AUG-2001; 2001WO-EP009892.  
 PR  
 XX (FARB ) BAYER AG.  
 PA  
 XX Tietjen K, Weidler M;  
 XX WPI; 2002-269010/31.  
 XX  
 XX Identifying plant target proteins for herbicidally active compounds,  
 PT comprising aligning and comparing nucleic acid or amino acid sequences  
 PT from plant with nucleic acid or amino acid sequences from non-plant  
 PT organisms.  
 XX  
 XX Claim 5; SEQ ID NO 896; 261pp + Sequence Listing; English.  
 XX  
 XX The invention relates to identifying target proteins (ABBS90790-ABB94016)  
 CC for herbicidally active compounds, comprising aligning and comparing  
 CC nucleic acid or amino acid sequences from plant with nucleic acid or  
 CC amino acid sequences from non-plant organisms using suitable search  
 CC parameters, where plant sequences having an E-value greater by a factor  
 CC of 3 than the E-value of most similar non-plant sequences are selected.  
 CC The polypeptides or nucleic acids encoding them are useful for  
 CC identifying modulators. The identified modulators are useful as  
 CC herbicides  
 XX  
 XX Sequence 616 AA;  
 SQ

Query Match 53.0%; Score 1542.5; DB 5; Length 616;  
 Best Local Similarity 53.5%; Pred. No. 1e-135;  
 Matches 349; Conservative 59; Mismatches 111; Indels 133; Gaps 17;

QY 1 MITALDLHYLTAVPLVYAMTAYSGVRWRIFTPDQCSGINRFVAFVLLSFHFTS 60  
 Db 1 MITWHDLYTLTAVPLVYAMTAYSGVRWRIFTPDQCSGINRFVAFVLLSFHFTS 60  
 QY 61 TNDPFAMNLEFLAADTLQKAVLALLASRGSLSPRALGLDWSITLSLTPLNTLVMG 120  
 Db 61 TNDPFAMNLEFLAADTLQKAVLALLASRGSLSPRALGLDWSITLSLTPLNTLVMG 117  
 QY 121 IPLLIRMGYGASAGTLMVQVVLQCIITWYTLMLFLFYRAARALVLDQPDGAASIVSF 180  
 Db 118 IPLLIRMGYGASAGTLMVQVVLQCIITWYTLMLFLFYRAARALVLDQPDGAASIVSF 175  
 QY 181 RVSDVSVSLARGD-VELEAEPDGVAGAGVSRGGDAGRVTVTKSTSRSEAAASHH 239  
 Db 176 KVSESVSVSLDGHDFLETDAEI-----GNDGKLHVTVRKSNASR-----R 214

QY 240 SQTMPRVSNLQVEIYSLQSSNPTPRGSSFNHADFFNIVGAAAKGGGGAAGD----- 293  
 Db 215 SLMTTPRPNLTGAEIYSLSS-----TPRGSNFNHSDFYSVMGFP-----GRLSNFGPADL 266  
 QY 294 -----EEKAGCGG--GGHSPQPOAVAVPAKR----- 319  
 Db 267 YSVQSRGPTPRPSNFENNAVKYGFYNTNTSSVPAAGSYPAFPNPFSTGTGVSTKPNKI 326  
 QY 320 -----KDLHMLVWSSSASVPSERAHVHFGAGADHADVLAKGAQA 360  
 Db 327 PKNQOQLOKDSKASHDAKELHMFVWSSSASPVSD-----VFGGAGDNV-----ATE 375  
 QY 361 YDEYGRDDY-----SSRTKNGSGGAD-----KGGPTLSKLSNSTAQLY 399  
 Db 376 QSEQAKETRMVYSDQPRKSNARGGDDIGGLDSGEGEREIEKATAGLNKMGNSTAELE 435  
 QY 400 PKDDGEGRAAAVAPPASVMTLILIMVWKLTIRNNTYSSLIGVYVSLVSYRWGTEMPA 459  
 Db 436 AAGDGGGNGGTMPPTSVMTLILIMVWKLTIRNNTYSSLIGLWLVYVYRWVAMPX 495  
 QY 460 IIAARSISILSDAGLGMFSLGLFVALQPKIIACGNSVATFAMAVRFTITGPAIMAVAGIA 519  
 Db 496 ILQSSISILSDAGLGMFSLGLFVALQPKIIACGNSVATFAMAVRFTITGPAIMAVAGIA 555  
 QY 520 VGLRGVLLHIAIIVQALPQGIIVPFVFAKEYGVHPDILSTA--YG-----PIT 564  
 Db 556 IGLHGDLLRIAIVQALPQGIIVPFVFAKEYGVHPDILSTA--YG-----PIT 607

RESULT 11  
 AAY44270  
 ID AAY44270 standard; protein; 647 AA.  
 AC AAY44270;  
 XX  
 DT 28-FEB-2000 (first entry)  
 DE A. thaliana Ethylene Insensitive Root (EIR1) allele - EIR1-S97G.  
 KW Root-specific auxin transport protein; Ethylene Insensitive Root;  
 KW gravitropism; auxin-based herbicide; plant root; transgenic plant;  
 KW auxin homeostasis; EIR1 allele; EIR1-S97G;  
 KW altered auxin-mediated response.  
 XX Arabidopsis thaliana.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FT Misc-difference 97  
 FT /note= "Wild type Ser substituted by Gly"  
 XX  
 XX WO9963092-A1.  
 XX  
 XX 09-DEC-1999.  
 XX  
 XX 03-JUN-1999; 99WO-US012277.  
 XX  
 XX 03-JUN-1998; 98US-0087789P.  
 PR  
 XX (WHED ) WHITEHEAD INST BIOMEDICAL RES.  
 PA  
 XX Luschning C, Gaxiola RA, Grisafi P, Fink GR;  
 XX WPI; 2000-086979/07.  
 DR  
 XX DNA encoding a root-specific auxin transport protein, used to develop  
 PT transgenic plants with increased resistance to herbicides.  
 XX  
 XX Claim 17; Page; 55pp; English.  
 CC The present sequence is an EIR1 allele which was obtained by replacing  
 CC the conserved Ser97 of A. thaliana EIR1 with Gly. EIR1 protein functions

as a root-specific auxin transport (efflux) carrier and is involved in gravitropism. The EIR1 allele confers altered auxin-mediated responses in plants. The allele can be used to produce transgenic plants with high sensitivity to fluorinated indolic compounds, and mutant plants in which the roots are agravitropic and have reduced sensitivity to ethylene. EIR1 sequence is used for producing genetically engineered plants which have greater resistance to auxin-based herbicides and auxin transport inhibitors, and exhibit altered auxin homeostasis. Note: The present sequence is not shown in the specification but has been derived from EIR1 protein sequence shown in figure-7

Query Match 53.0%; Score 1541; DB 3; Length 647;  
Best Local Similarity 51.6%; Pred. No. 1.5e-135;  
Matches 348; Conservative 64; Mismatches 115; Indels 148; Gaps 17;

QY 1 MITALDLYHVLTAFLVAVLAVLALL-----ALASGLSPRALGDSITLFSITLPT 116  
DB 1 MITGKMDYDLAAWPLVYAMILAYGSRVWRIPTDQCNGINRFVALFAPVLLSFHFI 60  
QY 61 TNDPFAMNLFALADTLQKAVLALL-----ALASGLSPRALGDSITLFSITLPT 113  
DB 61 SNDPYAMNHFLLAADSLOKVVILAAFLWQAFSRRG-----GLEWMTLFSITLPT 113  
QY 117 LVNGIPLLRGMYGASAGTLMVQVVLQCIITWYTLMLFLFEYRAARALVLDQFDGAAS 176  
DB 114 LVNGIPLLRGMYGASAGTLMVQVVLQCIITWYTLMLFLFEYRAARALVLDQFDGAAS 176  
QY 177 IVSFRVDSVDSL-ARGDVELEABPDGVAGAGVSSRGDAGRVVTVRKSTSRSEAAC 235  
DB 172 ITSFRVDSVISLNGREPLQDAEI-----GDDGLHVVVRRSSAASSMISS 218  
QY 236 ---SHS---HSQTMQPRVSNLGSVEIYSLOSRRNPTPRGSSFNHADFF----- 277  
DB 219 FNKSHGGGLNSMITPRASNLTGVEIYSVQSSREPTPRASSFNQDTFYAMFNASKAPSPR 278  
QY 278 ---NIVGAAAKGGG----- 288  
DB 279 HGYTNSYGGAGAGPGGDVYVLSQSKGVTPRTSNFDEEVMKTKAKAGRGRSMGELYNN 338  
QY 289 ---CAAGDEEKACGGGGGSHSPQPAVAVPAKRKDLHMLVWSSA 331  
DB 339 SVPSYPPNPMFTGSTSGASGVKKKESGGSGG-----GVGVGGQNKEMNFWSSA 392  
QY 332 SPVSERAAVHVFAGGA-----DHADVLAKGAQAYDEY---GRDYSSTRTKNG 376  
DB 377 SGADKGGPTLKLGSNSTAQLYPKDDGEGRAAAVAMPASVMTLLIIMVWRKLIENPN 436  
QY 453 NNGGK--SPYMKKGGSD-----VEDGGPGPRKQKQMPASVMTLLIIMVWRKLIENPN 503  
DB 437 TYSSLIGVWVSLVSRWGT:EMPAITARSISILSDAGLWAMFSLGLFWALQPRIIACNK 496  
DB 504 TYSSLIGVWVSLVSRWGT:EMPAITARSISILSDAGLWAMFSLGLFWALQPRIIACNK 496  
QY 497 LAIANGVRFVAGPAWMAAASVAVLGRVLLHIAIQAALPQGIIVFFFAKEYGVHPDIL 556  
DB 564 VAGFAMVRFVAGPAWMAAASVAVLGRVLLHIAIQAALPQGIIVFFFAKEYGVHPDIL 556  
QY 557 STA--YG-----PIT 564  
DB 624 STAVIEGLVALPVT 638

RESULT 12  
ID AAY53129  
ID AAY53129 standard; protein; 647 AA.  
XX  
AC AAY53129;  
XX  
DT 03-APR-2000 (first entry)

XX Gravitropism stimulation response related protein AGR.  
DE DE Gravitropism stimulation response; AGR; plant root.  
XX Gravitropism stimulation response; AGR; plant root.  
XX Arabidopsis thaliana.  
OS  
XX JP11318463-A.  
PN  
XX 24-NOV-1999.  
PD  
XX 15-MAY-1998; 98JP-00134097.  
PF  
XX 15-MAY-1998; 98JP-00134097.  
PR  
XX (KAZU-) ZH KAZUSA DNA KENKYUSHO.  
PA (OJIP) OJI PAPER CO.  
XX  
DR WPI; 2000-100768/09.  
DR N-PSDB; AAZ57348, AAZ57349.  
XX  
A gene participating to the gravitropism stimulation response of a plant root.  
PT  
XX  
PS Claim 1; Page 9-11; 20pp; Japanese.  
XX  
CC The present sequence represents a protein which participates in the gravitropism stimulation response of a plant root. The protein is designated AGR. AGR can be used to improve the fixing rate of a plant root to soil  
CC  
XX  
SQ Sequence 647 AA;

Query Match 52.7%; Score 1535; DB 3; Length 647;  
Best Local Similarity 51.4%; Pred. No. 5.6e-135;  
Matches 347; Conservative 64; Mismatches 116; Indels 148; Gaps 17;

QY 1 MITALDLYHVLTAFLVAVLAVLALL-----ALASGLSPRALGDSITLFSITLPT 116  
DB 1 MITGKMDYDLAAWPLVYAMILAYGSRVWRIPTDQCNGINRFVALFAPVLLSFHFI 60  
QY 61 TNDPFAMNLFALADTLQKAVLALL-----ALASGLSPRALGDSITLFSITLPT 116  
DB 61 SNDPYAMNHFLLAADSLOKVVILAAFLWQAFSRRG-----GLEWMTLFSITLPT 113  
QY 117 LVNGIPLLRGMYGASAGTLMVQVVLQCIITWYTLMLFLFEYRAARALVLDQFDGAAS 176  
DB 114 LVNGIPLLRGMYGASAGTLMVQVVLQCIITWYTLMLFLFEYRAARALVLDQFDGAAS 176  
QY 177 IVSFRVDSVDSL-ARGDVELEABPDGVAGAGVSSRGDAGRVVTVRKSTSRSEAAC 235  
DB 172 ITSFRVDSVISLNGREPLQDAEI-----GDDGLHVVVRRSSAASSMISS 218  
QY 236 ---SHS---HSQTMQPRVSNLGSVEIYSLOSRRNPTPRGSSFNHADFF----- 277  
DB 219 FNKSHGGGLNSMITPRASNLTGVEIYSVQSSREPTPRASSFNQDTFYAMFNASKAPSPR 278  
QY 278 ---NIVGAAAKGGG----- 288  
DB 279 HGYTNSYGGAGAGPGGDVYVLSQSKGVTPRTSNFDEEVMKTKAKAGRGRSMGELYNN 338  
QY 289 ---CAAGDEEKACGGGGGSHSPQPAVAVPAKRKDLHMLVWSSA 331  
DB 339 SVPSYPPNPMFTGSTSGASGVKKKESGGSGG-----GVGVGGQNKEMNFWSSA 392  
QY 332 SPVSERAAVHVFAGGA-----DHADVLAKGAQAYDEY---GRDYSSTRTKNG 376  
DB 377 SGADKGGPTLKLGSNSTAQLYPKDDGEGRAAAVAMPASVMTLLIIMVWRKLIENPN 436  
QY 453 NNGGK--SPYMKKGGSD-----VEDGGPGPRKQKQMPASVMTLLIIMVWRKLIENPN 503  
DB

QY 437 TYSSLGIVVMSLVSRWGIEMPAIARSISILSDAGLGMAMFSLGLFMALQPRIIACGNK 496  
Db 504 TYSSLFGLAWSLVSPKNIKMTIMSGSISILSDAGLGMAMFSLGLFMALQPKIACGKS 563  
QY 497 LAAIANGVRVAGAPVMAAASINVLGRVILHIAIQAALPQGIIVPFVFAKEYGVHPDIL 556  
Db 564 VAGFAMAVRFLTGPAVIAATSAIGIRGDLHLHIAIQAALPQGIIVPFVFAKEYGVHPDIL 623  
QY 557 STA--YG-----PIT 564  
Db 624 STAVIFGMLVALPVT 638

RESULT 13

AA44264  
ID AAY44264 standard; protein; 647 AA.

XX AC AAY44264;

XX DT 28-FEB-2000 (first entry)

XX DE A. thaliana Ethylene Insensitive Root (EIR1) protein.

XX KW Root-specific auxin transport protein; EIR1; Ethylene Insensitive Root;  
KW gravitropism; auxin-based herbicide; plant root; transgenic plant;  
XX auxin homeostasis; plasma membrane; transmembrane domain.

XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers  
FT Peptide 1..60  
FT Domain 10..28  
FT FT /label= Signal\_peptide  
FT FT /label= Transmembrane\_domain  
FT FT 41..60  
FT FT /label= Transmembrane\_domain  
FT FT 71..92  
FT FT /label= Transmembrane\_domain  
FT FT 100..121  
FT FT /label= Transmembrane\_domain  
FT FT 132..152  
FT FT /label= Transmembrane\_domain  
FT FT 503..520  
FT FT /label= Transmembrane\_domain  
FT FT 530..553  
FT FT /label= Transmembrane\_domain  
FT FT 568..589  
FT FT /label= Transmembrane\_domain  
FT FT 595..611  
FT FT /label= Transmembrane\_domain  
FT FT 622..642  
FT FT /label= Transmembrane\_domain

XX PN WO9963092-A1.

XX PD 09-DEC-1999.

XX PF 03-JUN-1999; 99WO-US012277.

XX PR 03-JUN-1998; 98US-0087789P.

XX PA (WHED ) WHITEHEAD INST BIOMEDICAL RES.

XX PI Luschnig C, Gaxiola RA, Grisafi P, Fink GR;

XX DR WPI; 2000-086979/07.

XX DR N-PSDB; AAZ29279, AAZ29278.

XX PT DNA encoding a root-specific auxin transport protein, used to develop

XX PT transgenic plants with increased resistance to herbicides.

XX PS Claim 18; Fig 7; 55pp; English.

XX XX

CC The present sequence is a EIR1 protein which functions as a root-specific  
CC auxin transport (efflux) carrier involved in gravitropism. The sequence  
CC was obtained from A. thaliana cDNA phage library. EIR1 consists of  
CC potential N-glycosylation sites and a N-terminal signal peptide which  
CC indicates localisation of the protein in the plasma membrane. The  
CC transmembrane domains are located in the highly conserved portions of the  
CC protein, N- and C-termini. The sequence is used for producing  
CC genetically engineered plants with greater resistance to auxin-based  
CC herbicides and auxin transport inhibitors in combination with a second  
CC herbicide. It can also be used to enhance transport of auxin in plant  
CC roots, produce transgenic plants which exhibit altered auxin homeostasis  
CC and mutant plants in which the roots are agravitropic and have reduced  
CC sensitivity to ethylene

XX Sequence 647 AA;

Query Match 52.7%; Score 1535; DB 3; Length 647;

Best Local Similarity 51.4%; Pred. No. 5 6e-135;

Matches 347; Conservative 64; Mismatches 116; Indels 148; Gaps 17;

QY 1 MITALDLHYHVTAVVPLIYAMTLAGSVRWIRIFTDQCSGINRFVALFAVLLSFHFTS 60  
Db 1 MITGKMDYDLAAMVPLIYAMTLAGSVRWIRIFTDQCSGINRFVALFAVLLSFHFTS 60  
QY 61 TNDPFAMNLRFLAADTLQKVAVLALL-----ALASRGLSSPRALGLDWSITLPSLSLTPNT 116  
Db 61 SNOPYANVYHFLAADSILQKVILALLFLWQAFSRG-----SLEWMTLISLTPNT 113  
QY 117 LVMGIPILLRGMYGASSAGITLVQVIVLQCIITLMLFLFEYRAARALVLDQPDGAAS 176  
Db 114 LVMGIPILLRGMYGASSAGITLVQVIVLQCIITLMLFLFEYRAARALVLDQPDGAAS 176  
QY 177 IVSPFRVDSVVSLL-ARGDVELEAEPDGVAGAGAVSRGSDAGRVVTVKASTSSSEAAAC 235  
Db 172 ITSFRVDSVVSLLNGREPLQTDAEI-----GDDGKLHVVRSSAASSMISS 218  
QY 236 ---SHS---HSQTMQPRVNSLGSVEIYSLQSSNPTPGSSFNHADFF----- 277  
Db 219 FNKSHGGLNLSMITPRASNLGTGVEIYVQSSREPTPRASSFNQTD FYAMFNASKAPSR 278  
QY 278 ---NIVGAAAKGGG----- 288  
Db 279 HGYTNSYGGAGAGPGDGVYLSQSSKGVTPRTSNFDEVMKTKAGRGSRWSGELYNN 338  
QY 289 -----GAAGDEKACGGGGGHSQPQAVAVPAKRDLMHVLWSSSA 331  
Db 339 SVPSYPPNPMTFTGSGASGVKKKSGGGSGG-----GVGVGQNKEMMFVWSSSA 392  
QY 332 SPVSERAAVHVFAGGA-----DHADVLAKGAQAYDEY---GRDDYSRTKNG 376  
Db 393 SPVSEANAKNMTRGSDVSTDPKVSIPPHDNLATKAMQNLINMSPGRGKHVEMDQDG 452  
QY 377 SGGADKGGPTLSKLGSTAGLYPKDDGEGRAAAVAMPASPATRLIILIMVWRKLIRPN 436  
Db 453 NNGGK--SPYMGKKGSD-----VEDGGFGRKQMPASPATRLIILIMVWRKLIRPN 503  
QY 437 TYSSLGIVVMSLVSRWGIEMPAIARSISILSDAGLGMAMFSLGLFMALQPRIIACGNK 496  
Db 504 TYSSLFGLAWSLVSPKNIKMTIMSGSISILSDAGLGMAMFSLGLFMALQPKIACGKS 563  
QY 497 LAAIANGVRVAGAPVMAAASINVLGRVILHIAIQAALPQGIIVPFVFAKEYGVHPDIL 556  
Db 564 VAGFAMAVRFLTGPAVIAATSAIGIRGDLHLHIAIQAALPQGIIVPFVFAKEYGVHPDIL 623  
QY 557 STA--YG-----PIT 564  
Db 624 STAVIFGMLVALPVT 638

RESULT 14

AA44271

ID AAY44271 standard; protein; 647 AA.

XX XX



CC produce a more robust root system, alter root angle or redirect root  
CC growth; and may be useful for gene mapping (e.g. for plant breeding) and  
CC to identify loss of function mutants. This protein may be useful for  
CC raising specific antibodies, for the detection of auxin transport  
CC proteins and to design and/or identify specific inhibitors of auxin  
CC transport proteins, potentially useful as herbicides  
XX

XX	Sequence 647 AA;
SQ	Query Match 52.7%; Score 1535; DB 3; Length 647;
	Best Local Similarity 51.4%; Pred. No. 5.6e-135;
	Matches 347; Conservative 64; Mismatches 116; Indels 148; Gaps 17;
QY	1 MITALDLHYLTAVVPLVYAVMTLAYGSVRWWEIFTDQCQGINRVALFAVPLLSFHFIS 60
DB	1 MITGKMDYDVLAAVPLVYAVMTLAYGSVRWWEIFTDQCQGINRVALFAVPLLSFHFIS 60
QY	61 TNDPFAMLRFLAADTLQKAVLALL-----ALASRGLSPRALGLDWSITLFSLTLPNT 116
DB	61 SNDPYAMNVHFLAADSLOKXVILAALFLWQAFPRRG-----SLEWMITLFSLTLPNT 113
QY	117 LVMGIPLLRGMTGASAGTLMVQVVLQCIITWYTLMLFLFEYRAARALVLDQPDGAAS 176
DB	114 LVMGIPLLRAMYGDFS-GNLMVQIVVLOSIIWYTLMLFLFEFRGAKLILISEQFPE-TAGS 171
QY	177 IVSFRVDSVWSL-ARGDVELEAEPDGVAGAGAVSSRGDAGRVRVTVRKSTSRSEAAC 235
DB	172 ITSFRVDSVVISUNGREPLQTDAEI-----GDDGKLHVVRSSAASSMISS 218
QY	236 ----SHS----HSQTMOPRVSNLSGVIIYSLOSRRNTPRGSSFNHADFF----- 277
DB	219 FNKSHGGGLNLSMITPRASNLTGVEIYSVQSSREPTPRASSFNQTDIFYAMFNASKAPSPR 278
QY	278 ----NIVGAAAKGGG----- 288
DB	279 HGVTNSYGGAGAGPGGDVYVYSSLOSXKGVTPRTSNPDEEVMTAKAGRGSRMSGELYNNN 338
QY	289 -----GAAGDEEXGACGGGGGHSPOQAVAPAKRKDLHMLVWSSSA 331
DB	339 SVPSYPPPPNPMFTGTSKGASGVKKESGGGSGG-----GVGVGGQNKEMMFVWSSSA 392
QY	332 SPVSEBAAVHVFAGGA-----DHADVLAKGAQAYDEY---GRDDYSSTKNG 376
DB	393 SPVSEANAKNAMTRGSSSTVDTPKVISIPPHDNLATKAMQNLINNSPGRKHVEMDQDG 452
QY	377 SGGADKGGPTLSKLGNSSTAQLYPKDDGGRAAAAMPPASVMTRLIILIMVWRKLIRNPN 436
DB	453 NNGGK--SPYMGKKGSD-----VEDGGPGPKQMPPEASVMTRLIILIMVWRKLIRNPN 503
QY	437 TVSSSLGVVWSLVSVYRWGTEMPAIIARSISILSDAGLGNAMFSLGLFMALQPRIIACGNK 496
DB	504 TVSSSLFGLAWSLVSPKWNIMKPTIMSGSISILSDAGLGNAMFSLGLFMALQPKIICGKS 563
QY	497 LAAIAMGVRFVAGPVMMAASIAVGLRGVLLHIAIVQAAALPQGIVPFVFAKEYGVHPDIL 556
DB	564 VAGFANAVRFLTGPAVIAATSAIGIRGDLHLHIAIVQAAALPQGIVPFVFAKEYNVHPDIL 623
QY	557 STA--YG-----PIT 564
DB	624 STAVIFGLVALPVT 638

Search completed: March 3, 2004, 09:07:00  
Job time : 80 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: March 3, 2004, 09:09:24 ; Search time 608 Seconds  
(without alignments)  
4003.647 Million cell updates/sec

Title: US-10-030-884-14  
Perfect score: 2910  
Sequence: 1 MIALDLYHVTAVVPIYA.....DILSTAVGPTSHGFITCHS 573

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Dgapop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool/US10030884/runat\_24022004\_103922\_1940/app\_query.fasta\_1.711  
-DB=N\_Geneseq\_29Jan04 -QMT=fascap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-USER=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOXT=7

- Database : N\_Geneseq\_29Jan04:\*
- 1: Geneseqn1980as:\*
  - 2: Geneseqn1990as:\*
  - 3: Geneseqn2000as:\*
  - 4: Geneseqn2001as:\*
  - 5: Geneseqn2001bs:\*
  - 6: Geneseqn2002as:\*
  - 7: Geneseqn2003as:\*
  - 8: Geneseqn2003bs:\*
  - 9: Geneseqn2003cs:\*
  - 10: Geneseqn2004as:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2910	100.0	2769	AAA94721	AAA94721 Corn auxi
2	2129.5	73.2	1926	ADA70391	Ada70391 Rice gene
3	2020.5	69.4	1686	ADA69470	Ada69470 Rice gene
4	1665	57.2	2374	AAZ29280	AAZ29280 Rice EIR1
5	1647.5	56.6	1827	ADA70668	Ada70668 Rice gene
6	1643	56.5	2549	AAA94729	AAA94729 Soybean a
7	1641.5	56.4	2292	AAV16361	AAV16361 cDNA sequ
8	1596	54.8	1845	ADA69396	Ada69396 Rice gene

9	1592	54.7	1884	7	ADA70233	Ada70233 Rice gene
10	1566.5	53.8	2293	3	AAA94733	AAA94733 Wheat aux
11	1560.5	53.6	1923	3	AAZ44722	AAZ44722 Arabidops
12	1560.5	53.6	1923	6	ABZ13042	ABZ13042 Arabidops
13	1542.5	53.0	1851	3	AAZ43229	AAZ43229 Arabidops
14	1535	52.7	2204	3	AAZ29279	AAZ29279 A. thalia
15	1535	52.7	2232	3	AAZ57348	AAZ57348 Gravitrop
16	1527	52.5	2324	6	AAZ94731	AAZ94731 Soybean a
17	1526	52.4	1860	6	ABZ13043	ABZ13043 Arabidops
18	1526	52.4	1860	7	ADA68467	Ada68467 Arabidops
19	1526	52.4	2027	3	AAZ44892	AAZ44892 Arabidops
20	1513	52.0	2162	3	AAA94724	AAA94724 Rice auxi
21	1408.5	48.4	1725	3	AAZ46518	AAZ46518 Arabidops
22	1406.5	48.3	1983	6	ABZ13500	ABZ13500 Arabidops
23	1338	46.0	2117	3	AAV16362	AAV16362 cDNA sequ
24	1318.5	45.3	2061	3	AAZ46177	AAZ46177 Arabidops
25	1299.5	44.7	2101	3	AAZ94727	AAZ94727 Soybean a
26	1228	42.2	7072	3	AAZ29278	AAZ29278 A. thalia
27	1218.5	41.9	3980	3	AAZ57349	AAZ57349 Gravitrop
28	1183	40.7	1713	3	AAZ48532	AAZ48532 Arabidops
29	1004	34.5	1426	3	AAA94719	AAA94719 Corn auxi
30	965	33.2	1618	3	AAA94725	AAA94725 Soybean a
31	662.5	22.8	1088	3	AAA94716	AAA94716 Corn auxi
32	630	21.6	855	3	AAA94737	AAA94737 Corn auxi
33	454	15.6	504	3	AAA94720	AAA94720 Corn auxi
34	385	13.2	425	3	AAA94736	AAA94736 Wheat aux
35	383.5	13.2	415	3	AAA94735	AAA94735 Wheat aux
36	356	12.2	473	3	AAA94732	AAA94732 Wheat aux
37	334	11.5	330	3	AAA94723	AAA94723 Rice auxi
38	323	11.1	531	3	AAA94726	AAA94726 Soybean a
39	322	11.1	419	3	AAA94730	AAA94730 Soybean a
40	320.5	11.0	543	3	AAA94722	AAA94722 Rice auxi
41	316	10.9	624	3	AAA94718	AAA94718 Corn auxi
42	303.5	10.4	629	3	AAA94715	AAA94715 Corn auxi
43	273	9.4	525	3	AAA94728	AAA94728 Soybean a
44	259.5	8.9	447	3	AAA94734	AAA94734 Wheat aux
45	257	8.8	620	7	ABZ73099	ABZ73099 Rice leaf

ALIGNMENTS

RESULT 1  
AAA94721  
ID AAA94721 standard; DNA; 2769 BP.  
XX  
AC AAA94721;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Corn auxin transport protein clone p0119.cmtnl24r DNA sequence.  
XX  
KW Auxin transport protein; corn; root development; gene mapping;  
XX  
OS plant breeding; herbicide; ss.  
XX  
OS Zea mays.  
XX  
PN WO200068389-A2.  
XX  
PD 16-NOV-2000.  
XX  
PF 03-MAY-2000; 2000WO-US012061.  
XX  
PR 07-MAY-1999; 99US-0133040P.  
XX  
PA (DUPO) DU PONT DE NEMOURS & CO E I.  
XX  
PI (PTON-) PIONEER HI-BRED INT INC.  
XX  
PI Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
XX  
DR WPI; 2000-687647/67.  
XX  
DR P-PSDB; AAB26934.  
XX  
PT New nucleic acid sequences encoding new auxin transport proteins, useful

PT for modulating root growth of plants and to screen for herbicides.

PS Claim 2; Page 60-61; 9app; English.

XX Auxins are plant hormones that influence plant behaviour and development  
XX e.g. vascular tissue differentiation, apical development, tropic  
XX responses and organ (e.g. flower, leaf) development. The present  
XX invention relates to corn auxin transport protein coding sequences and  
XX proteins. The present sequence is one such coding sequence. This sequence  
XX may be used to modulate root development, e.g. to produce a more robust  
XX root system, alter root angle or redirect root growth. Also, the present  
XX sequence may be useful for gene mapping (e.g. for plant breeding) and to  
XX identify loss of function mutants. The protein encoded by the present  
XX sequence may be useful for raising specific antibodies, for the detection  
XX of auxin transport proteins and to design and/or identify specific  
XX inhibitors of auxin transport proteins, potentially useful as herbicides  
XX  
SQ Sequence 2769 BP; 527 A; 917 C; 757 G; 568 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.41e-194 Length: 2769  
Score: 2910.00 Matches: 573  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-10-030-884-14 (1-573) x AAA94721 (1-2769)

QY 1 MetlleThrAlaLeuAspLeuTyHisValLeuThrAlaValValProLeuTyValAla 20  
DB 171 ATGATCACCGGGTGGAGCTTACACAGTGTGTACGCGGGTGTGGCGTGTACGTGGCC 230  
QY 21 MetThrLeuAlaTyGlySerValArgTrpArgilePheThrProAspGlnCysSer 40  
DB 231 ATGAGCTGGGACGGTCCGTCGGTGTGGGGCAGTCTTACGCGGACAGTGTCTCC 290  
QY 41 GlylleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
DB 291 GGGATCAACCGCTTTCGGTGGCTTTCGGCGCTTCGGCGTTCCTTCCATCTTCATCTCC 350  
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
DB 351 ACCAAGACCCCTTCGCATGAACTCGGCTTCTGGCCGCCACACAGCTGCGAGAGTG 410  
QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
DB 411 GCCGTCTTCGGGCTGTGGCGCTTCCGCGCTTCCTTCCCGCGCGCGCTCGGG 470  
QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
DB 471 CTGGACTGGAGCATCACGCTTCTTCTCTCTCCACGCTTCCCAACAGCTGTGTATGGGC 530  
QY 121 IleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
DB 531 ATCCCGCTGTCTGGAGGCATGTACGGCGCTGTGGCGCGCTTCTTCCCGCGCGCGTCCAGGTC 590  
QY 141 ValValLeuGlnCysIleIleTrpTyThrLeuMetLeuPheLeuPheGluTyArgAla 160  
DB 591 GTGCTCTCCAGTGCATCACTGTGTACAGCTCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 650  
QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
DB 651 GCGCGCGGCTCGTCTCTGACAGTTCCTCCGACGGCGCGCGCTTCTTCTTCTTCTTCTTCTTCTTCT 710  
QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
DB 711 CGGCTGACTCCGAGCTGCTCTGCTCTGCCAGGGGGAGCTCGAGCTCGAGCGCGAGGCC 770  
QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
DB 771 GACGGCGTCCGCGCGCGCGCGCTCTCTTCCCGCGCGCGGACCGCGCGGGTGGCC 830

QY 221 ValThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSer 240  
DB 831 GTCACTGTGGCAAGTCCACAGCTCGCTCGAGCCCGGTGCTCGCACTCGCACTCC 890  
QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTySerLeuGlnSer 260  
DB 891 CAGACCATGAGCCCGCTGTGTCTCAACTCTCGCGGTGGAGATCTACTCGTGCAGTCG 950  
QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheIleVal 280  
DB 951 TCGCGCAACCCACCCCGCGCGGTTCAGCTTCAACAGCCGACTTCTTCAACATCGTC 1010  
QY 281 GlyAlaAlaLysGlyGlyGlyAlaAlaGlyAspGluGlyGlyAlaCysGly 300  
DB 1011 GCGCGCGCGCGCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1070  
QY 301 GlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLys 320  
DB 1071 GCGCGCGCGCGAGGACACTCGCGCGCGCGCGCGCTGCGCTGCGCGCGCGCGAGAG 1130  
QY 321 AspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaVal 340  
DB 1131 GACCTGCACATGCTGTGTGGAGCTTCAGCGCTTCGCGCGTGTTCGAGCGCGCGCGCG 1190  
QY 341 HisValPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAla 360  
DB 1191 CACGTCTTCGGCGCGCGCGCTGACCATGCGAGCTCTCGCCAAAGAGCGCGCGCG 1250  
QY 361 TyrAspGluTyGlyArgAspAspTySerSerArgThrIleAsnGlySerGlyAla 380  
DB 1251 TACGACAGTACCGCGCGCGAGCTACAGCAGCAGGACGAAAGACGCGCGCGCGCG 1310  
QY 381 AspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyPro 400  
DB 1311 GACAAAGCGCGCGCGCGCTGTGAGAGCTGGGGTCCAACTCGACGCGCGCGAGTATCC 1370  
QY 401 LysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThr 420  
DB 1371 AAGGACACCGCGCGCGCGCGCGCGGTGGCGATCCCGCGCGCGAGCTGTATGACG 1430  
QY 421 ArgLeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTySerSer 440  
DB 1431 CGGCTCATCTCTCATCATGTGTGGAGAGAGCTGATCCGGAAACCCCAACACCTTACTCCAGC 1490  
QY 441 LeuIleGlyValValTrpSerLeuValSerTyArgTrpGlyIleGluMetProIleIle 460  
DB 1491 CTCATCGCGCTCTGTGTCTGTGTCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1550  
QY 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480  
DB 1551 ATGCGCGCGTTCGATTCGATTCCTGTGACGCGCGGTCTCGGGATGGCCATGTTCAGCCTA 1610  
QY 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIle 500  
DB 1611 GGCTGTTCATGCGCTGCGCGCGCGAGGATCATCGCGTGGCGGAAACAAGCTGCGCGCATC 1670  
QY 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaVal 520  
DB 1671 GCGATGGCGCTCGGTTCGTTCGACAGCCCGCGGTTCATGGCGCGCGCTTCATCGCGCTC 1730  
QY 521 GlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIle 540  
DB 1731 GGTGTGCGCGCGCTCTCTCTCCATCGCATCGCTGTCCAGGCTGTCTGTCTCAGGGGATC 1790  
QY 541 ValProPheValPheAlaLysGluTyGlyValHisProAspIleLeuSerThrAlaTy 560  
DB 1791 GTGCGCTTCGTGTTCGCCAAGGATACGGGTTCATCCCGACATCTCTGACGACAGCGTAT 1850  
QY 561 GlyProIleThrSerHisGlyPheIleThrCysHisSer 573  
DB 1851 GGTCCAATAACATCGATGTTTCATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1889

RESULT 2



QY 419 MetThrArgLeuLeuLeuMetValTrpArgLysLeuLeuArgAsnProAsnThrTyr 438  
 Db 1186 ATGACGAGGCTCATCTGATCATGTTTGGAGGAGCTGATCAGGAACCCCAACTTAC 1245  
 QY 439 SerSerLeuLeuGlyValValTrpSerLeuValSerTyrArgTyrGlyLeuMetPro 458  
 Db 1246 TCAGGCTCTTGGTGTATCTGTCATCTGTCCTCCCTGCTCTACAGTGGGGATTGAGTGCCA 1305  
 QY 459 AlaIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPhe 478  
 Db 1306 GCATCATCGCCGGTTCGATTTCGATCTCTGATGCAGATGCAGGCTTGGATGGCCATGTC 1365  
 QY 479 SerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAla 498  
 Db 1366 AGCTAGGATTTGTCATGTCATTCGACGACGATCATTTGCTGTGGGAATCTCCCTGCT 1425  
 QY 499 AlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIle 518  
 Db 1426 TCGATGTCATGCGCTGAGTTCCTCGTGGTCCGCGCTGATGCTGCCCTCCATC 1485  
 QY 519 AlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleVal 532  
 Db 1486 GCGTCCGACTTCGCGGGTGTCTTCTGCACATTCGCAATTTGTCAGGTAGCTGCAAGTTC 1545  
 QY 532 532 532  
 Db 1546 AGGCTTCTGCACCAAAATGATCTTAACTGCAAAAGATCTATCGTAATCAACAATTG 1605  
 QY 533 533 533  
 Db 1606 TCATTGAATTCCTCAGGCGCTCTTCTCAAGGAATCGTCCCTTTGTGTTGCCAAGGAG 1665  
 QY 549 TyrGlyValHisProAspIleLeuSerThrAlaTyrGlyProIleThrSerHisGlyPhe 568  
 Db 1666 TACAATGTTTCATCTTAACATTCGTGAGCAGCGATGATCTCTTACCATCTTTCTTTT 1725

RESULT 3  
 ID ADA69470 standard; DNA; 1686 BP.  
 XX ADA69470;  
 AC ADA69470;  
 XX 20-NOV-2003 (first entry)  
 DT Rice gene, SEQ ID 2793.  
 XX Plant; bacterial infection; fungal infection; viral infection; rice;  
 XX gene; ds.  
 XX Oryza sativa.  
 XX WO2003000898-A1.  
 XX 03-JAN-2003.  
 XX 22-JUN-2001; 2001WO-IB001105.  
 XX 22-JUN-2001; 2001WO-IB001105.  
 XX (SYN) SYNGENTA PARTICIPATIONS AG.  
 XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
 PI Katagiri F, Qian S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
 XX WPI; 2003-175290/17.  
 XX Identifying at least one gene involved in plant resistance or response to  
 PT pathogenic infection for conferring resistance or tolerance to a plant to  
 PT bacterial, fungal or viral infection by determining or detecting plant  
 PT gene expression.  
 XX Claim 6; SEQ ID NO 2793; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes  
 CC involved in plant resistance or response to pathogenic infection. M1  
 CC comprises identifying a gene whose expression is significantly altered in  
 CC the incompatible interaction of plant gene expression relative to  
 CC expression of the gene in an uninfected plant, in a mutant plant that  
 CC does not express a gene associated with response to pathogenic infection,  
 CC or in a corresponding incompatible or compatible interaction. (M1) is  
 CC useful for conferring resistance to resistance or tolerance to a plant to  
 CC bacterial, fungal or viral infection. The present sequence was used to  
 CC illustrate the invention.  
 XX SQ Sequence 1686 BP; 286 A; 553 C; 490 G; 356 T; 0 U; 1 Other;  
 Alignment Scores:  
 Pred. No.: 1.81e-132 Length: 1686  
 Score: 2020.50 Matches: 442  
 Percent Similarity: 75.69% Conservative: 22  
 Best Local Similarity: 72.10% Mismatches: 39  
 Query Match: 69.43% Indels: 111  
 DB: 7 Gaps: 16  
 US-10-030-884-14 (1-573) x ADA69470 (1-1686)  
 QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 Db 1 ATGATCAGGTGGTGGACCTGTACACGCTCTGACGGCGGTGTGCGTGTGTAGTGGCG 60  
 QY 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
 Db 61 ATGACGTGGGTACGCGCTCCGTCGGTGGTGGGCACTTCCTCCCGACCACTGTCTCC 120  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisIleSer 60  
 Db 121 GGCATCAACCGCTTCGCGCTCTTCGCGTCCGCTCTCTCTCTCCACTTCATCTCC 180  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 181 ACCAACCAACCCCTTCGCCCATCACTCCGCTCTCTCCCGCCGACACGCTCCAGAACCTC 240  
 QY 81 AlaValLeuAlaLeuLeuAla 96  
 Db 241 ATCGTCTCGCCCTCTCTCGCGCTCTGTCGGTCTCTCGCGCGCGG----- 288  
 QY 97 ArgAlaLeuGlyLeuAspTrpSerIleThrIlePheSerLeuSerThrLeuProAsnTh 116  
 Db 289 -----TCCCTCGACTGGCTCATCACCTCTTCTCCCTCTCCAACTTCCCAAMAC 339  
 QY 116 rLeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLe 136  
 Db 340 CCTGTATGGGCATCCGCTGCTCAAGGGAT----- 372  
 QY 136 uMetValGlnValValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPh 156  
 Db 373 -----TGCATCATCTGGTACACGCTCATGCTGTCTCTT 407  
 QY 156 eGluTyrArgAlaAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaLase 176  
 Db 408 CGAGTACCGTGGCGCGGCTTCTTGTCTATGGAGCAGTTCCTCGGAC---ACCGCGGCTC 464  
 QY 176 rIleValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAsp----- 193  
 Db 465 CATGCTCTCTTCGCGGTGGATTCCGACGCTGCTCTGCTCGCGCGCGCGGTGGAGC 524  
 QY 194 -ValGluLeuGluAlaGluProAspGlyValAlaGlyAlaValSerSerArgGly 213  
 Db 525 CGCGGAGCTGCAGCGGAG-----GCGGAGGT 551  
 QY 213 yGlyAspAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAl 233  
 Db 552 CGGGGACGATGGCAAGATCGGAGTCACTCGTGGCGCAAGTCGACGAGCTCGGCTCCGAGGC 611  
 QY 233 alaCysSerHis-----SerHisSerGlnThrMetGlnProArgValSerAsnLe 250

Db	612	GGCGTGTCTCGCAGCGGACGAGTACACACTCGAGTCCATGCGCGCGCTCCCAACCT	671
Qy	250	userGlyValGluIleuTyrSerLeuGlnSerArgAsnProThrProArgGlySerSe	270
Db	672	CTCCGGGTGAGATTACTCGCTGCGAGTCTCGCGAACCAGCGCGCGGGCTCCAG	731
Qy	270	rPheAsnHisAlaAspPhePheAsnIleValGlyAlaAlaAlaGlyGlyGlyAl	290
Db	732	CTTCAACACACCGCGAGTCTTCAACATCGTC	773
Qy	290	aAlaGlyAspGluGluGlyAlaCyGlyGlyGlyGlyGlyGlyHisSerProGlnPr	310
Db	774	GCACGGGACAGAGAGAGCGGC-----GCCGGTGGCGGGGCGCACATCGCGCAGCC	827
Qy	310	oGlnAlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSe	330
Db	828	G-----GTGGTGGCGAAGAGAGGACCTGCACATGTCGTGTGGAGCTCAAG	875
Qy	330	rAlaSerProValSerGluArgAla-----AlaValHisValPh	343
Db	876	CGCTCGCGGTGTCGAGCGCGCGCGCGCGCGCTCGCGCGCGTGCACGTCCT	935
Qy	343	eGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTyrAspG	363
Db	936	CGGCGTGTGGCGCGCCGACCGCGAC-----GCCAAGGTGCTCAGCGCTATGATGA	989
Qy	363	uTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysG	383
Db	990	G-----TACAGCTTCGGGAACAGAAAT-----GAGAAGGA	1019
Qy	383	yGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAs	403
Db	1020	CGGGCCGACACTGTCGAAGCTGGGGTCCACTCGACGGCGGAGCTCCGGCCAAAGGACGA	1079
Qy	403	pGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThrArgLeuI	423
Db	1080	CGGCGAGGGAGGGCGCA-----GCGATCGCGCGCGAGCGTGATGACGAGGCTCAT	1133
Qy	423	eLeulleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleG	443
Db	1134	ACTGATCATGTCTGGAGAGAGTGCATCAGAACCCAAACACTTACTCCAGCCCTCCGG	1193
Qy	443	yValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIleIleAlaAr	463
Db	1194	TGTCACTGTCCTCTCTCTACAGTGGGGATTGAGATCGCGCTATCATCGCCCG	1253
Qy	463	gSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPh	483
Db	1254	GTGATTTTCGATCTCTTCAGATGCGGGCTCGGAATGCCCATGTTCAGCCCTAGGATTGT	1313
Qy	483	eMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMetG	503
Db	1314	CATGGCATTCAGCCACCGATCATTCCTGTGGGACTCCCTTGTCTGATGCAATGGC	1373
Qy	503	yValArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuAr	523
Db	1374	CGTCAGGTTCTCGTGTGTCCTCGCTCATGGCTCGCGCTCCATCGCGCTCGGACTTCG	1433
Qy	523	gGlyValLeuLeuHisIleAlaIleVal-----	532
Db	1434	CGGGTGTCTTCGCACATTGCCATTGTTTCAGTAAGTGCAGTTCAGGCTTCGACCA	1493
Qy	533	-----	533
Db	1494	AAATGCATCTTAACCTGCAAGAAATCTATCGTAACTCAACAAATGTCATTGAACTCTCA	1553
Qy	533	nAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlnTyrGlyValHisPr	553
Db	1554	GGCCGCTCTTCCTCAAGGAATCGTCCCTTCGTGTTCGCAAGGAGTACATATGTTATCC	1613
Qy	553	oAspIleLeuSerThrAlaTyrGlyProIleThrSer	565

Db	1614	TAAcATTCTGAGCACAGCGTATGATCTCTTTACCACC	1650
RESULT	4		
AAZ29280			
ID	AAZ29280	standard; cDNA; 2374 BP.	
XX			
AC	AAZ29280;		
XX			
DT	28-FEB-2000	(first entry)	
XX			
DE	Rice EIR1 Homologue (REH1) cDNA.		
XX			
KW	Root-specific auxin transport protein; EIR1; Ethylene Insensitive Root;		
KW	gravitropism; auxin-based herbicide; plant root; transgenic plant;		
KW	auxin homeostasis; Rice EIR1 Homologue; REH1; ds.		
XX			
OS	Oryza sativa.		
XX			
PH	Key	Location/Qualifiers	
FT	CDS	158..1945	
FT		/*tag= a	
FT		/product= "REH1 protein"	
FT	sig_peptide	158..337	
FT		/*tag= b	
XX			
PN	WO9963092-A1.		
XX			
PD	09-DEC-1999.		
XX			
PF	03-JUN-1999;	99WO-US012277.	
XX			
PR	03-JUN-1998;	98US-0087789P.	
XX			
PA	(WHED ) WHITEHEAD INST BIOMEDICAL RES.		
XX			
PI	Luschnig C, Gaxiola RA, Grisafi P, Fink GR;		
XX			
DR	WPI; 2000-086979/07.		
DR	P-PSDB; AAV44265.		
XX			
PT	DNA encoding a root-specific auxin transport protein, used to develop		
PT	transgenic plants with increased resistance to herbicides.		
XX			
PS	Claim 1; Fig 8; 55pp; English.		
XX			
CC	The present sequence is a cDNA encoding REH1, a rice homologue of EIR1		
CC	which functions as a root-specific auxin transport (efflux) carrier		
CC	protein involved in gravitropism. The sequence is obtained from a rice		
CC	EST derived from root-specific cDNA. The sequence is used for producing		
CC	genetically engineered plants with greater resistance to auxin-based		
CC	herbicides and auxin transport inhibitors in combination with a second		
CC	herbicide. It can also be used to enhance transport of auxin in plant		
CC	roots, produce transgenic plants which exhibit altered auxin homeostasis		
CC	and mutant plants in which the roots are agravitropic and have reduced		
CC	sensitivity to ethylene		
XX			
SQ	Sequence 2374 BP; 471 A; 736 C; 733 G; 434 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	2,26e-107	Length:	2374
Score:	1665.00	Matches:	365
Percent Similarity:	67.58%	Conservative:	54
Best Local Similarity:	58.87%	Mismatches:	111
Query Match:	57.22%	Indels:	90
DB:	3	Gaps:	17
US-10-030-884-14	(1-573) x AAZ29280	(1-2374)	
Qy	1	MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla	20
Db	158	ATGATTACGCGCGCGGACTTCACCGATGACGCGCGGTCGCTGACGTGGCG	217
Qy	21	MetThrLeuAlaTyrGlySerValArgTrpArgIlePheThrProAspGlnCysSer	40





```

1576 GCGTGTCTCCATGGCCCTCGCTTCTTCGGGGCCCTGGCGTATGGCGCGGTCFCA 1635
518 IIEAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuPro 537
1636 ATCGCATCGGACTCGCGGGACGCTCTGCGACGTCGCCATTGTCAGCGGCTCTACCA 1695
538 GlnGlyIleValProPheValPheAlaLysGlyIleValHisProAspIleLeuSer 557
1696 CAAGGATGTGCTCTTTGTTTGGCAAGAATAACAATGTCCACCGGCCATCCTGAGC 1755
558 ThrAla-----TyrGly-----ProIleThr 564
1756 ACAGCGTAAATTTTGGCATGCTAATAGCTCTTCCATCACA 1797

```

RESULT 6  
 ID AAA94729 standard; DNA; 2549 BP.  
 AC AAA94729;  
 DT 02-FEB-2001 (first entry)  
 DE Soybean auxin transport protein clone sf11.pk131.g9 DNA sequence.  
 KW Auxin transport protein; soybean; root development; gene mapping;  
 KW plant breeding; herbicide; ss.  
 OS Glycine max.  
 PN WO200068389-A2.  
 PD 16-NOV-2000.  
 PF 03-MAY-2000; 2000WO-US012061.  
 PR 07-MAY-1999; 99US-0133040P.  
 PA (DUPO) DU PONT DE NEMOURS & CO E I.  
 PA (PION-) PIONEER HI-BRED INT INC.  
 PI Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
 XX WPI; 2000-687647/67.  
 DR P-PSDB; AAB26942.  
 DR New nucleic acid sequences encoding new auxin transport proteins, useful  
 PT for modulating root growth of plants and to screen for herbicides.  
 PS Claim 2; Page 74-75; 94pp; English.

XX Auxins are plant hormones that influence plant behaviour and development  
 CC e.g. vascular tissue differentiation, apical development, tropic  
 CC responses and organ (e.g. flower, leaf) development. The present  
 CC invention relates to corn auxin transport protein coding sequences and  
 CC proteins. The present sequence is one such coding sequence. This sequence  
 CC may be used to modulate root development, e.g. to produce a more robust  
 CC root system, alter root angle or redirect root growth. Also, the present  
 CC sequence may be useful for gene mapping (e.g. for plant breeding) and to  
 CC identify loss of function mutants. The protein encoded by the present  
 CC sequence may be useful for raising specific antibodies, for the detection  
 CC of auxin transport proteins and to design and/or identify specific  
 CC inhibitors of auxin transport proteins, potentially useful as herbicides  
 XX  
 SQ Sequence 2549 BP; 736 A; 572 C; 518 G; 723 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 8,55e-106 Length: 2549  
 Score: 1643.00 Matches: 351  
 Percent Similarity: 66.51% Conservative: 58  
 Best Local Similarity: 57.30% Mismatches: 111  
 Query Match: 56.46% Indels: 100  
 3 Gaps: 17

```

US-10-030-884-14 (1-573) x AAA94729 (1-2549)
Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
Db 309 ATGATCACCTTACAGACITCTACCATGTGATGCAATGGTGGCCATCTATGTGGCC 368
Qy 21 MetThrIleuAlaTyrGlySerValArgTyrPArgIlePheThrProAspGlnCysSer 40
Db 369 ATGATACTAGCCCTATGGCTCAGTGAAGTGGTGAAGATTTTCTCCCTGATCAATGCTCT 428
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
Db 429 GGCATCAACCGTTTGTGGACCTCTTGGCAGTGCCTCTCTCTCTCTCCATTCATAGCC 488
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
Db 489 TCCACACAAACCTTATGAGATGAACCTGAGGTTCCTAGCTGCTGACACCTTCAAAAGATC 548
Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96
Db 549 ATAATACTAGTCTCTCTTGGCAGTTTGGAGCAACATCACCAAAAGGGTGT----- 599
Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116
Db 600 -----TTGGAATGGGCATAAACCTTGTCTCTCTCCACCTTCCCAACACT 647
Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136
Db 648 TTGTTATGCGATCCCTTTGCTCAAGGAGATATGGTACTTCTCA-----GGAGGCTC 704
Qy 137 MetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156
Db 705 ATGGTGCAAAATTTGGTCTCTCCAGTCATCATTTGGTACACCTTGATGCTCTTCTTGT 764
Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaSer 176
Db 765 GAGTTTAGAGTGCCAGAAATGCTCATCTCTGACAGTTCCTCTGAC---ACTGTGCTCTCC 821
Qy 177 IleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGlu 195
Db 822 ATTGTCTCCATCCATGTGGACTCTGATGTCATGTCATGGATGGAAGACCAACACTTGAG 881
Qy 196 LeuGluAlaGluProAspGlyValAlaGlyAlaValSerSerArgGlyGlyAsp 215
Db 882 ACTGAAGCTGAGATCAAG-----GAA 902
Qy 216 AlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys 235
Db 903 GATGGTAAACTCCATGTCATCTGTGAGGAATCCAATGCTTCAAGATCAGACATCTTCTCT 962
Qy 236 SerHisSerHisSer---GlnThrMetGlnProArgValSerAsnLeuSerGlyValGlu 254
Db 963 AGAAGGTCTCAGGGTCTCTCTTCCACCATCCACGCCCTTCCAACTTACCAATGCTGAG 1022
Qy 255 IleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAla 274
Db 1023 ATATCTCTTTGCAATCTCTAGAACCTTACCGAGAGGCTCTAGTTTCAACCACT 1082
Qy 275 AspPhePheAsnIleValGlyAlaAlaLysGlyGlyGlyAlaAlaGlyAsp--- 293
Db 1083 GATTTCTACTCCATGATGGTCTGCTGGTGGCAGGAACCTCAAACTTTGGTCTGATGTT 1142
Qy 294 -----GluGluLysGly 297
Db 1143 TATGGCCTTTTCAAGTTCAGAGGGCCCACTCCAGGCCCTTCTAATGATGATGATGGT 1202
Qy 298 Ala-----CysGlyGlyGlyGlyGlyGlyHisSerProGlnProGln 311
Db 1203 GGGAAAGCCAAAGTTTTCATTACCATGCTGCTGGTGGAACTGGGCACTACCTCGCACCAAC 1262
Qy 312 -----AlaValAlaValProAlaLys 318

```

Db	1263	CTGGCATGTTCTCTCCCTCTAATGGGTCCAAAAGTGTGTGCTGAATGTAATGCCAAG	1322	
Qy	319	Arg	-----LysAspLeuHisMetLeu	325
Db	1323	AGGCCTAATGGCAGGCTCAGCTGAAGCCTGAGGATGGGAATAGGACCTTCATATGTTT	1382	
Qy	326	ValTrpSerSerSerAlaSerProValSerGluArgAlaValHisValPheGlyAla	345	
Db	1383	GTTTGGAGTTCAAGTGTTCACCAAGTTTCTGATGTGTTGGTCCCATGAGTAGTGA--	1439	
Qy	346	GlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnLaTyrAspGluTyrGly	365	
Db	1440	GGAGGTGATCATCAGAAAGAAGTCAAATTCGATCTCCAGGAAAGTGGAGAATAAT	1499	
Qy	366	-----ArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAsp	381	
Db	1500	CATAGACACTCAAGAAGACTACTTAGAGAAAGATGATTCAGCTTTGGGAATAGAGAA	1559	
Qy	382	LysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLys	401	
Db	1560	ATG-----GACAGGGAGATGAATCAGCTTGAGGTGAG	1592	
Qy	402	AspAspGlyGluGlyArgAlaAlaValAlaMetProProLaserValMetThrArg	421	
Db	1593	AAGGTTGGAGATGGAAACCAAAA-----ACCATGGCTCCAGCAAGTGTGATGACAAAG	1646	
Qy	422	LeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeu	441	
Db	1647	CTTATATGATATGGTGTGGAGAAACTCATCAGAAACCCCAACCTACTCTAGCCTA	1706	
Qy	442	IleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIleIle	461	
Db	1707	ATTGGTCTCACTTGGTCTCTTTGTTTCATTCAAGTGGAAATGTTGAGATGCCCTGCCAATA	1766	
Qy	462	AlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGly	481	
Db	1767	GCAAGGCTATCTCCATATTGTCCAGACGACAGGCTTGGCATGGCCATGTTTCAGTCTTGGT	1826	
Qy	482	LeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAla	501	
Db	1827	CTCTTCATGCTTTCACCCAGGGTCATAGCATGTGGAAATTCACAGCAGCTTTTGCC	1886	
Qy	502	MetGlyValArgPheValAlaGlyProAlaValaMetAlaAlaSerIleAlaValGly	521	
Db	1887	ATGGCTGTGAGATTCCCTTACAGGTCACAGCTGTGTCATGGCAGCTGCTTCCATTGCTTGGG	1946	
Qy	522	LeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleVal	541	
Db	1947	CTCAAGGTGTTCTCCCTACAGTTGCCATTGTTTCAGGCAGCTCTTCCCAGAGAAATTGC	2006	
Qy	542	ProPheValPheLysGlyTyrGlyValHisProAspIleLeuSerThrAla-----	559	
Db	2007	CCATTGCTCTTGTAAAGGAATATATGTACATCTCTGATATCTCAGCAGAGCTGTATT	2066	
Qy	560	TyrGly-----ProIleThr	564	
Db	2067	TTTGGGATGCTGATGCTTTGCCCATAACT	2096	
RESULT	7			
ID	AAV16361			
XX	AAV16361	standard; cDNA to mRNA; 2292 BP.		
AC	AAV16361;			
XX	25-MAR-2003	(revised)		
DT	08-JUN-1998	(first entry)		
XX				
XX	cDNA sequence of the pin1 gene (pcpin1/23) of Arabidopsis thaliana.			
XX	pin1; pcpin1/23; auxin polar transport; auxin receptor; auxin carrier;			
KW	auxin efflux protein; pin homologue; auxin transport; alteration; growth;			
KW	wood production; morphogenesis; vascular tissue differentiation;			
KW	abscission; fruit ripening; ds.			





533	Qy	GlnAlaIalaLeuProGlnGlyIleValProPheValPheAlaLysGlyTyrGlyValHis	552
1699	Db	CAGGAGAGTCTGCGCCCAAGGCGATTGCTTCGTCAGGAGTACAGGTCGAC	1758
553	Qy	ProAspIleLeuSerThrAla-----TyrGly-----ProIleThr	564
1759	Db	CCTAGCATTCACAGCAGGTGTCATCTTTGGCATGCTCATCGCTTCGCTATCACC	1815

## RESULT 9

ADA70233

ID ADA70233 standard; DNA; 1884 BP.

XX

AC ADA70233;

XX 5

DT . 20-NOV-2003 (first entry)

DE  
PR  
CO  
CO

DE Rice gene, SE(
xx

Plant: bacterial

KW : gene: ds

25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858

OS *Oryza sativa*.

XX  
XX

PN WO2003000898-A1.

XX

PD 03-JAN-2003.

[illegible]

PF 22-JUN-2001; 2001WO-IB001105.

[illegible]PR 22-JUN-2001; 20  
YYXX  
PA  
(SYCN ) SYNGENTA PARTICIPATIONS AG

PA (SIGN ) SINGENTA PARTICIPATIONS AG.  
XX

PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y:

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX  
XX

DR WPI; 2003-175290/17.



PT Identifying at least one gene involved in plant resistance or response to

pathogenic infection for conferring resistance or tolerance to a plant to

bacterial, fungal or viral infection by determining or detecting plant

PT gene expression.

XX  
PS CLAIM 6. SEO TD NO 3556. PROJ: 60

XX  
PS CLAIM 6; SEQ

CC The present invention relates to a method (M1) for id

involved in plant resistance or response to the present invention includes CO and methods involved in plant resistance or response to

comprises identifying a gene whose expression

CC the incompatible interaction of plant gene

CC expression of the gene in an uninfected pla

CC does not express a gene associated with re

or in a corresponding incompatible or

useful for conferring resistance or tolerance to a plant to

CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the infection

CC  
yy  
illustrate the invention.

Db 1054 AAGGAGGTCGATGGCGTCCGCTCGCGCGCAAGCGGACGGCGTGGAGAGGACGAC 1113  
Qy 369 TyrSerSer-----ArgThrLysAsnGlySerGlyGly-AlaAspLysGlyG1 384  
Db 1114 TTCAGCTT-CGGGAACAGAGCGCTCGCGAGAGGACGCGGAGCGCGGACGAGAGAG 1172  
Qy 384 yProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspG1 404  
Db 1173 CGTGGCGCGCGGTGTCGGGTGAGCATGGCAAGCTGGCTTGACGCG----- 1221  
Qy 404 yGluGlyArgAlaAlaAlaValMetProProAlaSerValMetThrArgLeuLeuLe 424  
Db 1222 -----GGCGCGAGCGGATGCCCGCAGCGCGTGTGATGACGGCGCTCATCCT 1268  
Qy 424 uileMetValTyrArgLysLeuLeuArgAsnProAsnThrTyrSerSerLeuLeuGlyVa 444  
Db 1269 GATCATGTGTGGCGACAGCATCCCGCAACCCCAACACCTACTCCAGCCTCATCGGCT 1328  
Qy 444 lValTyrSerLeuValSerTyrArgTyrGlyLeuMetProAlaLeuLeuLeuAlaArgSe 464  
Db 1329 CATCTGGTGGTCTGCTGCTGCTGCGTGAACCTTCAGATGCGCGCATCATCTGAATC 1388  
Qy 464 rLleSerLleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMe 484  
Db 1389 CATCTCATCTCTCCGACGCGCGCTCGGATGGCCATGTTGAGCTCGGCGCTGTTTCAT 1448  
Qy 484 tAlaLeuGlnProArgLleLeuAlaCysGlyAsnLysLeuAlaAlaLeuAlaMetGlyVa 504  
Db 1449 GCGCTGCGAGCGCGATCATCGCTGCGGGAACAAGGTGGCGACGTTCCGCGATGGCGT 1508  
Qy 504 lArgPheValAlaGlyProAlaValMetAlaAlaSerLleAlaValGlyLeuArgG1 524  
Db 1509 GCGGTCTCTGACCGCGCGCGCGCTCATGGCGCGCGCTCCATCGCGCGCTCGCGCG 1568  
Qy 524 yValLeuLeuHisLleAlaLleValGlnAlaAlaLeuProGlnGlyLleValProPheVa 544  
Db 1569 CACGCTCTCCAGCTGCCATCGTCGACGAGCGGTTCGCMCAGCGCATTTGTCCTTSGT 1628  
Qy 544 lPheAlaLysGluTyrGlyValHisProAspLleLeuSerThrAla 559  
Db 1629 CTTGCCAAGGAGTACAGCTGCACCCCGACATTTCTACGACGCGG 1674

RESULT 10  
AA94733  
ID AAA94733 standard; DNA; 2293 BP.

XX AC AAA94733;  
XX AC  
XX DT  
XX DE 02-FEB-2001 (first entry)  
XX DE Wheat auxin transport protein clone wdkic.pk008.g1 DNA sequence.  
XX DE Auxin transport protein; wheat; root development; gene mapping;  
XX KW Plant breeding; herbicide; ss.  
XX OS Triticum aestivum.  
XX XX WO200068389-A2.  
XX XX 16-NOV-2000.  
XX XX 03-MAY-2000; 2000WC-US012061.  
XX XX 07-MAY-1999; 99US-0133040P.  
XX XX (DUPO) DU PONT DE NEMOURS & CO E I.  
XX XX (PION-) PIONEER HI-BRED INT INC.  
XX XX Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
XX WPI: 2000-687647/67.  
XX DR P-PSDB; AAB26946.  
XX XX

PT New nucleic acid sequences encoding new auxin transport proteins, useful  
PT for modulating root growth of plants and to screen for herbicides.  
XX Claim 2; Page 84; 94pp; English.  
XX Auxins are plant hormones that influence plant behaviour and development  
CC e.g. vascular tissue differentiation, apical development, tropic  
CC responses and organ (e.g. flower, leaf) development. The present  
CC invention relates to corn auxin transport protein coding sequences and  
CC proteins. The present sequence is one such coding sequence. This sequence  
CC may be used to modulate root development, e.g. to produce a more robust  
CC root system, alter root angle or redirect root growth. Also, the present  
CC sequence may be useful for gene mapping (e.g. for plant breeding) and to  
CC identify loss of function mutants. The protein encoded by the present  
CC sequence may be useful for raising specific antibodies, for the detection  
CC of auxin transport proteins and to design and/or identify specific  
CC inhibitors of auxin transport proteins, potentially useful as herbicides  
XX

SQ Sequence 2293 BP; 462 A; 713 C; 683 G; 435 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 1.75e-100 Length: 2293  
Score: 1566.50 Matches: 364  
Percent Similarity: 65.26% Conservative: 53  
Best Local Similarity: 56.96% Mismatches: 132  
Query Match: 53.83% Indels: 91  
DB: 3 Gaps: 17

US-10-030-884-14 (1-573) x AAA94733 (1-2293)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
Db 74 ATGATCACCGGGAGAGACATCTACGACGTCGTGGCGCGTGGTGGCTGTACGTGGCC 133  
Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTyrArgIlePheThrProAspGlnCysSer 40  
Db 134 ATGTTTCATGCGGTACGGTTCGGTTCGGTGGCGCATCTTCACGCGGACCATGCTCG 193  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 194 GGATCAACCGCTTCGTGCGCGTCTTCGGCGTGGCGCTCTCTCTCTCTCTCTCTCTCT 253  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
Db 254 ACCACGACCCCTACGCCATGGACTACCGCTTCTTGGCGCGGACTCGCTGAGAGAGTC 313  
Qy 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
Db 314 GTCATCTCGCGCCCTCGCGGTGGGACAAACGTCGTCTCCGCTACCGGTGCGCGCGC 373  
Qy 101 -----LeuAspTyrSerIleThrLeuPheSerLeuSer 111  
Db 374 GGACGAGGCGCGCGAGGCTCGTGGTGGAGTGGACCATCACGCTCTTCTCTCTGGCG 433  
Qy 112 ThrLeuProAsnThrLeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSer 131  
Db 434 ACGTGTCCCAACACGCTGGTGTGGCATCCGCTCTGCGCGCCATGTACGCGGACTTC 493  
Qy 132 SerAlaGlyThrLeuMetValGlnValValValLeuGlnCysIleIleTyrThrLeu 151  
Db 494 TCG---GGGTGCGTTCATGGTGCAGATCTGGTGTGTCGAGCGGTCTCTGGTACAGCTC 550  
Qy 152 MetLeuPheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAsp 171  
Db 551 ATGCTCTTCTCTCGAGTACCGCGCGCGGACAGCGCTCATCTCCGAGCATTTCCGCGCC 610  
Qy 172 GlyAlaAlaAlaSerIleValSerPheArgValAspSerAspValValSerLeuAlaArg 191  
Db 611 GACGTGGCGCGCGAGCATCGCTCTTTCGGGTGCATCTCGAGCTCTCTCTCTCTCTCT 667  
Qy 192 GlyAspValGluLeuGluAlaGluProAspGlyValAlaAlaGlyAlaGlyAlaValSer 211  
Db 668 GGGCGCGAGCGCTGCACGCGGAC-----GCC 694

513	Qy	513	talaalaasertlealavalglyLeuArgGlyValleuLeuHislealalleValGI	533
:	:	:		:
1768	Db	1768	CGCGCGACCTCAATCGCGCTCGGGGAGTCTCTACATGTTGCCATGTGCCA	1827
:	:	:		:
533	Qy	533	nalaaleuProGInglyIleValProPheValPheAlalysGluTyrGlyValHisPr	553
:	:	:		:
1828	Db	1828	GGCAGCATTTCCACAGGAATTTGTTCCATTTGTTTCCCAAGGAGTACAATTGCCATCC	1887
:	:	:		:
553	Qy	553	asPilleuSerThrAla-----TyrGly-----ProlleThr	564
:	:	:		:
1888	Db	1888	TCAAATACTATGACACAGCGGTTATTTTGGAAATCTCGTGGCGCTCCCGATCAG	1942
:	:	:		:
RESULT 11				
AAC44722				
ID	AAC44722 standard; DNA, 1923 BP.			
XX	AAC44722;			
DT	18-OCT-2000 (first entry)			
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 43888.			
XX	Arabidopsis thaliana.			
OS	Arabidopsis thaliana.			
XX	EP1033405-A2.			
PN	XX			
PD	XX			
PP	06-SEP-2000.			
PF	XX			
FF	25-FEB-2000; 2000EP-00301439.			
XX	XX			
XX	25-FEB-1999; 99US-0121825P.			
PR	05-MAR-1999; 99US-0123180P.			
PR	09-MAR-1999; 99US-0123548P.			
PR	23-MAR-1999; 99US-0125788P.			
PR	25-MAR-1999; 99US-0126264P.			
PR	29-MAR-1999; 99US-0126785P.			
PR	01-APR-1999; 99US-0127452P.			
PR	06-APR-1999; 99US-0128234P.			
PR	08-APR-1999; 99US-0128714P.			
PR	16-APR-1999; 99US-0129845P.			
PR	19-APR-1999; 99US-0130077P.			
PR	21-APR-1999; 99US-0130449P.			
PR	23-APR-1999; 99US-0130510P.			
PR	23-APR-1999; 99US-0130891P.			
PR	28-APR-1999; 99US-0131449P.			
PR	30-APR-1999; 99US-0132048P.			
PR	30-APR-1999; 99US-0132407P.			
PR	04-MAY-1999; 99US-0132484P.			
PR	05-MAY-1999; 99US-0132485P.			
PR	06-MAY-1999; 99US-0132486P.			
PR	06-MAY-1999; 99US-0132487P.			
PR	07-MAY-1999; 99US-0132863P.			
PR	11-MAY-1999; 99US-0134256P.			
PR	14-MAY-1999; 99US-0134218P.			
PR	14-MAY-1999; 99US-0134219P.			
PR	14-MAY-1999; 99US-0134221P.			
PR	18-MAY-1999; 99US-0134370P.			
PR	19-MAY-1999; 99US-0134769P.			
PR	20-MAY-1999; 99US-0135124P.			
PR	21-MAY-1999; 99US-0135353P.			
PR	24-MAY-1999; 99US-0135629P.			
PR	25-MAY-1999; 99US-0136021P.			
PR	27-MAY-1999; 99US-0136392P.			
PR	28-MAY-1999; 99US-0136782P.			
PR	01-JUN-1999; 99US-0137222P.			
PR	03-JUN-1999; 99US-0137528P.			
PR	04-JUN-1999; 99US-0137502P.			
PR	07-JUN-1999; 99US-0137724P.			

Hybridisation assay; Genetic mapping; gene expression control;  
protein identification; signal transduction pathway; metabolic pathway;  
promoter; termination sequence; ss.

PR	08-JUN-1999;	99US-0138094P.	PR	09-AUG-1999;	99US-0147493P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147935P.
PR	10-JUN-1999;	99US-0138847P.	PR	10-AUG-1999;	99US-0148171P.
PR	14-JUN-1999;	99US-0139111P.	PR	11-AUG-1999;	99US-0148319P.
PR	16-JUN-1999;	99US-0139452P.	PR	12-AUG-1999;	99US-0148341P.
PR	16-JUN-1999;	99US-0139453P.	PR	13-AUG-1999;	99US-0148565P.
PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	15-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	22-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	22-JUN-1999;	99US-0139899P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140353P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	18-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155132P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0156599P.
PR	09-JUL-1999;	99US-0142920P.	PR	25-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159330P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159638P.
PR	20-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0144814P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	21-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160815P.
PR	22-JUL-1999;	99US-0145192P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160989P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0161404P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145918P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161360P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146386P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161922P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161933P.
PR	03-AUG-1999;	99US-0147038P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147204P.			
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			

Alignment Scores:  
Pred. No.: 3,75e-100  
Score: 1560.50  
Percent Similarity: 63.87%

Length:  
Matches: 1923  
Conservative: 353  
73

Best Local Similarity:	52.92%	Mismatches:	102
Query Match:	53.63%	Indels:	139
DB:	3	Gaps:	20
US-10-030-884-14 (1-573) x AAC44722 (1-1923)			
QY	1	MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla	20
DB	1	ATGATCTCAGGACACACCTCTACACGGTCCCTACCGCGGTGATTCCTCTCTACGTGGCC	60
QY	21	MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer	40
DB	61	ATGATCTCGCTTACGGCTCTGTCGGTGTGGAAATCTTCTCACCGACCAATGCTCC	120
QY	41	GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer	60
DB	121	GGAAATCAACCGTTTGTTCGCCATCTTTCGGCGTCCCTCTCTCTCTCTCTCTCTCTCT	180
QY	61	ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal	80
DB	181	ACCACAACTCTTACGCAATGATCTCCGGTTCATCGCGCGGACACTCTCCAAAATC	240
QY	81	AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly	100
DB	241	ATCATGTGTCTCTAGTCTCTATGGCT-----AATTTCATCGCTCCCGGTAGC	291
QY	101	LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly	120
DB	292	CTGAGTGGAGCATCAATCTTTCCTCTCCACACTTCCACACTCTTCTGTATGGG	351
QY	121	IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal	140
DB	352	ATTCTCTCTTGATCGCATGTATGGCAATACTCT---GGTTCCCTCATGTGCCAAATC	408
QY	141	ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAla	160
DB	409	GTCTCTCTCAGGTATCATCTGTGTACACCTTCTCTCTTCTCTCTCTCTCTCTCTCT	468
QY	161	AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe	180
DB	469	GCCAAGATGCTCATCTAGGAGCAGTTCCTGTAG---ACGCTGCTTCCATTTCTTTCT	525
QY	181	ArgValAspSerAspValValSerLeuAlaArgGlyAsp---ValGlnLeuGluAlaGlu	199
DB	526	AAAGTCGAATCCGACGCTGTTCTGCTGACGGCCATGATTTCTTGAGACCGATGCGAG	585
QY	200	ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal	219
DB	586	ATA-----GGTGACGACGCGGAAGCTT	606
QY	220	ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis	239
DB	607	CACGTCAACCGTGAGAAATCAACGCTTACGT---CGTTGCTTCTGC-----	651
QY	240	SerGlnThrMetGlnProAspGlnValSerAsnLeuSerGlyValGluIleTyrSerLeuGln	259
DB	652	GGCCCGAACATGATCTCACGGCGCTCAAACTCACCGAGCTGAGATTTATAGTCTCAGC	711
QY	260	SerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIle	279
DB	712	ACC-----ACTCTAGAGGCTCTAAATTCAACCACTCTGATTTTTCACATG	759
QY	280	ValGly-----Ala	282
DB	760	ATGGGTTCCTCCGGTGTCTCTCTCCAAATTCGGTCCGGCGGATGATGACTCCGTCAA	819
QY	283	AlaAlaLysGly-----GlyGlyGlyAlaAlaGly	292
DB	820	TCATCTAGAGGTCCAACTCTCGACCTTCAAACTTCGAGGAGAAATTCGGCCATGGCATCC	879
QY	293	AspGluGlnLysGlyAlaCysGlyGlyGlyGlyHisSerProGlnPro-----	310
DB	880	TCCCGAGATTCGGGTATTATCCCTCGAGGAGGACCGCGGTCTTATCCGGCTCCGAATCCA	939
QY	310	-----	310
DB	940	GAGTTCCTCTCAACCAACCATCTACCGCAATAAAGCGTCAATAAACCAGAACAC	999
QY	311	-----GlnAlaValAlaValProAla-----LysArg	319
DB	1000	GTATAATACGAATCAGCAGACGACTCTTCCAAACGGGGCAAGTCAACACGCCATGACGCC	1059
QY	320	LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAla	339
DB	1060	AAGAGCTTCACATGTTCTGAGCTCAACCGGTTCACCGCTTTCAGACCGGGCGGT	1119
QY	340	ValHisValPheGlyAlaGlyGlyAlaAspHis-----AlaAspValLeu	354
DB	1120	CTTAACGTTTTCGTC---GGAGCACCTGTACAACGATCAAGCGCGAAGATCTGCCAAGGT	1176
QY	355	AlaLys-----	356
DB	1177	GCTAAAGAGATCCGTATGTAGTCCAGATCAATCTCAACCGCGAGACCAAGCTGTA	1236
QY	357	-----GlyAlaGlnAlaTyrAspGluTyrGlyArgAsp	367
DB	1237	GCTCATCCACCAAGTGGAGATTTTCGGAGGAGAACCAATTTAGTTTCGGCGGAAAGAA	1296
QY	368	AspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeu	387
DB	1297	GAAGAAGCAGACAGACCAAAAGACCGCGAGATGGT-----CTA	1335
QY	388	SerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArg	407
DB	1336	AACAACTTGTCTCAATTCACGGCGCGCTACAATCCAAG---ACAGGTCTAGGAGGA	1392
QY	408	AlaAlaAlaValAla-----MetProAlaSerValMetThrArgLeuIleLeu	424
DB	1393	GCCGAAGCAAGTCAACGAAAAATATGCTCCGCGAGTGTGTATGACAGAGCTATATCTG	1452
QY	425	IleMetValTrpArgGlyLeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyVal	444
DB	1453	ATAATGGTTTGGAGAACTCATCAGAAACCCAAACACTTACTTACTTCTCTCTCTCTCT	1512
QY	445	ValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIleIleAlaArgSer	464
DB	1513	ATTGGGCTCTCGTCCGTTTCCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC	1572
QY	465	IleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMet	484
DB	1573	ATCTCCATCTATCTGATGCTGGTCTTGGAAATGCAATGTTAGTTTGGGGTCTCTCATG	1632
QY	485	AlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyVal	504
DB	1633	CGGTTCACACCAATTAATTCGTTTGGGAATTCAGTGGCAACGTTTGGCATGCGCGTT	1692
QY	505	ArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuArgGly	524
DB	1693	AGGTTCCTTACGGGTCCGCGGTGATGCGGTTCGCTATAGCCATCGGATTACGTGGT	1752
QY	525	ValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheVal	544
DB	1753	GATTTACTCGGTGCTGCTATAGTTCAGGCCGCTTACCTCAAGGAATGTGCCCTTTGTG	1812
QY	545	PheAlaLysGlyTyrGlyValHisProAspIleLeuSerThrAla-----TyrGly---	561
DB	1813	TTTGGAAAGGAGTACAATGTTTCATCTCTCTATTTTAAGTACAGGGGTAAATTTGGAAATG	1872
QY	562	-----ProIleThr	564
DB	1873	CTTATAGCGCTTCCGATCAG	1893
RESULT 12			
ABZ13042			
ID ABZ13042 standard; DNA; 1923 BP.			
XX			



Db 1237 GCTCATCCAGCAAGTGGAGATTTCGGAGGAGAACCAATTTAGTTTCGCCGGAAGAA 1296  
Qy 368 AspTyrSerSerArgThrThyAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeu 387  
Db 1297 GAGGAAGCAGAGACACCAAGACCCGAGATGGT-----CTA 1335  
Qy 388 SerTysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArg 407  
Db 1336 AACAACTTGTCTCCAAATTCACGCGCGGTGTACAATCCAAG--ACAGGTCTAGGAGGA 1392  
Qy 408 AlaAlaAlaValAla-----MetProProlAserValMetThrArgLeuIleLeu 424  
Db 1393 GCCGAAGCAAGTCAACGAAAAATATGCTCCGGGAGTGTGATGACAAAGCTGATATG 1452  
Qy 425 IleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyVal 444  
Db 1453 ATAATGGTTTGGAGGAACATCATCAGAACCCCAACACTTACTCTAGTCTCATTTGGACTT 1512  
Qy 445 ValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIleAlaArgSer 464  
Db 1513 APTTGGGTCTCGTCTGCTTCCGGTGGCAGCGTGGCAATGCCCAAAATCATTCACAAATCT 1572  
Qy 465 IleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMet 484  
Db 1573 ATCTCATCTATCTGATGCTGGTCTTGGATGGCAATGTTAGTTGGGGTGTTCATG 1632  
Qy 485 AlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMetGlyVal 504  
Db 1633 GCGTTGCAACCAATTAATCCGCTTGTGGGAATTCAGTGGCAACGTTTGGCATGCGCGTT 1692  
Qy 505 ArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuArgGly 524  
Db 1693 AGTTCCTTACGGTCCGCGCGTGTATGGCGTGTGCTGTATAGCATCGGATTACGTGGT 1752  
Qy 525 ValLeuLeuHisIleAlaIleValGlnAlaLeuProGlnGlyIleValProPheVal 544  
Db 1753 GATTTACTGCGTGTGCTATAGTTTACGCGCGCATTTACCTCAAGGAATTTGCCCCTTTGTG 1812  
Qy 545 PheAlaLysGlyTyrGlyValHisProAspIleLeuSerThrAla-----TyrGly--- 561  
Db 1813 TTTCGAGAGGTACAAATGTTCTATCTCTGCTATTTAAGTACAGGGGTAAATTTGGAAATG 1872  
Qy 562 -----ProIleThr 564  
Db 1873 CTTATAGCGTTCGATCAG 1893  
RESULT 13  
AAC43229  
ID AAC43229 standard; DNA; 1851 BP.  
XX AC AAC43229;  
XX DT 17-OCT-2000 (first entry)  
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 38501.  
XX KW Hybridisation assay; genetic mapping; gene expression control;  
XX KW protein identification; signal transduction pathway; metabolic pathway;  
XX KW promoter; termination sequence; ss.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX PD 06-SEP-2000.  
XX PF 25-FEB-2000; 2000EP-00301439.  
XX QY 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 05-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134319P.  
PR 14-MAY-1999; 99US-0134421P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.

PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144325P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.	Alignment Scores:		
PR	05-AUG-1999;	99US-0147302P.	Pred. No.:		
PR	05-AUG-1999;	99US-0147260P.	Score:		
PR	06-AUG-1999;	99US-0147303P.	Length:		
PR	06-AUG-1999;	99US-0147416P.	Matches:		
PR	09-AUG-1999;	99US-0147493P.	Score:		
PR	09-AUG-1999;	99US-0147935P.	Percent Similarity:		
PR	11-AUG-1999;	99US-0148171P.	Best Local Similarity:		
PR	12-AUG-1999;	99US-0148319P.	Query Match:		
PR	13-AUG-1999;	99US-0148341P.	DB:		
PR	13-AUG-1999;	99US-0148565P.	US-10-030-884-14 (1-573) x AAC43229 (1-1851)		
PR	13-AUG-1999;	99US-0148684P.	QY		
PR	16-AUG-1999;	99US-0149368P.	Db		
PR	17-AUG-1999;	99US-0149175P.	QY		
PR	18-AUG-1999;	99US-0149426P.	Db		
PR	20-AUG-1999;	99US-0149722P.	QY		
PR	20-AUG-1999;	99US-0149723P.	Db		
PR	20-AUG-1999;	99US-0149929P.	QY		
PR	23-AUG-1999;	99US-0149902P.	Db		
PR	23-AUG-1999;	99US-0149930P.	QY		
PR	25-AUG-1999;	99US-0150566P.	Db		
PR	26-AUG-1999;	99US-0150884P.	QY		
PR	27-AUG-1999;	99US-0151065P.	Db		
PR	27-AUG-1999;	99US-0151066P.	QY		
PR	27-AUG-1999;	99US-0151080P.	Db		
PR	30-AUG-1999;	99US-0151303P.	QY		
PR	31-AUG-1999;	99US-0151438P.	Db		
PR	01-SEP-1999;	99US-0151930P.	QY		
PR	07-SEP-1999;	99US-0152363P.	Db		
PR	10-SEP-1999;	99US-0153070P.	QY		
PR	13-SEP-1999;	99US-0153758P.	Db		
PR	15-SEP-1999;	99US-0154018P.	QY		
PR	16-SEP-1999;	99US-0154039P.	Db		
PR	20-SEP-1999;	99US-0154779P.	QY		
PR	22-SEP-1999;	99US-0155139P.	Db		
PR	23-SEP-1999;	99US-0155486P.	QY		
PR	24-SEP-1999;	99US-0155659P.	Db		
PR	28-SEP-1999;	99US-0156458P.	QY		
PR	29-SEP-1999;	99US-0156598P.	Db		
PR	04-OCT-1999;	99US-0157117P.	QY		
PR	05-OCT-1999;	99US-0157533P.	Db		
PR	06-OCT-1999;	99US-0157865P.	QY		
PR	07-OCT-1999;	99US-0158029P.	Db		



XX DNA encoding a root-specific auxin transport protein, used to develop  
 PT transgenic plants with increased resistance to herbicides.  
 XX  
 PS Claim 1; Fig 6; 55pp; English.

XX The present sequence is a cDNA encoding BIR1 which functions as a root-  
 CC specific auxin transport (efflux) carrier protein involved in  
 CC gravitropism. The sequence was obtained from A. thaliana cDNA phage  
 CC library. The sequence is used for producing genetically engineered plants  
 CC with greater resistance to auxin-based herbicides and auxin transport  
 CC inhibitors in combination with a second herbicide. It can also be used to  
 CC enhance transport of auxin in plant roots, produce transgenic plants  
 CC which exhibit altered auxin homeostasis and mutant plants in which the  
 CC roots are agravitropic and have reduced sensitivity to ethylene  
 XX

SQ Sequence 2204 BP; 592 A; 484 C; 541 G; 587 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.: 2,69e-98 Length: 2204  
 Score: 1535.00 Matches: 347  
 Percent Similarity: 60.89% Conservative: 64  
 Best Local Similarity: 51.41% Mismatches: 116  
 Query Match: 52.75% Indels: 148  
 DB: 3 Gaps: 17

US-10-030-884-14 (1-573) x AA229279 (1-2204)

QY 1 MetIleThAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 DB 19 ATGATCACCAGGCAAGACATGATGATGTTTACGGCTATGTCGCCCTATACGTTGCT 78  
 QY 21 MetThrLeuAlaTyrGlySerValArgTTPArgIlePheThrProAspGlnCysSer 40  
 DB 79 ATGATATTAGCCTATGCTCGGTAGCGTGTGGGGGATTTACACCGGACCAATGTTCC 138  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 DB 139 GGTATAAACCGGTTCGTTCGGTTCGGGGTTCCTCTCTCTCTTCATCTCC 198  
 QY 61 ThrAsnAspPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIleVal 80  
 DB 199 TCCATGATCCCTACGCAATGAATACCACTTCCCTCGCTGCTGATCTCTTTCAGAAAATC 258  
 QY 81 AlaValLeuAlaLeuLeu-----AlaLeuAlaSerArgGlyLeuSerSerPro 96  
 DB 259 GTTATCTCGCGCACTCTTTCTTTGGCAGGGGTTTAGCCGACAGGA----- 306  
 QY 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 DB 307 -----AGCCTAGAATGGATGATAACGCTCTTTTCACTATCAACACTGCTCTTAACAGC 357  
 QY 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
 DB 358 TTGGTAATGGGAATCCATTGCTTAGGGCGATGTACGGAGACTTCTCC---GGTAACCTA 414  
 QY 137 MetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156  
 DB 415 ATGGTCAGATCGTGGTGTTCAGAGCATCATATGTTATATGTTATGTTCTCTCTCTGTTT 474  
 QY 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
 DB 475 GAGTTCGGTGGGCTAAGCTTCTATCTCCGAGCAGTTCCTCCGAG---ACGGCTGGTTCA 531  
 QY 177 IleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGlu 195  
 DB 532 ATTACTTCTCTCAGAGTTGACTCTGATGTTATCTCTTAATGCCCGTGAACCCCTCCAG 591  
 QY 196 LeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAsp 215  
 DB 592 ACCGATGCCGAGATA-----GGAGAC 612  
 QY 216 AlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys 235

DB 613 GACGGAAGCTACACGTGTGTGTGGAAGATCAAGTGGCGCTCATCATGATCTCTTCA 672  
 QY 236 -----SerHisSer-----HisSerGlnThrMetGlnProArgValSerAsn 249  
 DB 673 TTCACCAATCTCACGGCGAGAGATTAACTCTCTCATGATAACCGCGAGCTTCAAT 732  
 QY 250 LeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySer 269  
 DB 733 CTCACCGGGGTAGAGATTCTTACTCCGTTCAATCGTCACGAGAGCCGACCGAGAGCTTCT 792  
 QY 270 SerPheAsnHisAlaAspPhePhe----- 277  
 DB 793 AGCTTTAATCAGACAGATTCTTACGCAATGTTTAACGCAAGCAAAAGTCCCAAGCCCTCGT 852  
 QY 278 -----AsnIleValGlyAlaAlaLysGlyGly----- 288  
 DB 853 CACGGTTACACTAATAGCTACGGCGCGCTGAGCTGCTCCAGGTGGAGATGTTTACTCA 912  
 QY 288 ----- 288  
 DB 913 CTTCACTCTTCTAAAGCGGTGACGCGGAGAACGTCAAATTTTGTAGGAAGTTATGAAG 972  
 QY 288 ----- 288  
 DB 973 ACGGCGAAGAAAGCAGGAGAGGAGGAGATGATGAGTGGGGAATTATACACCAATAAT 1032  
 QY 289 -----GlyAlaAla 291  
 DB 1033 AGTGTTCGTCTACCCACCGCGGACCCCAATGTTACGGGGTCAACGAGTGGAGCAAGT 1092  
 QY 292 GlyAspGluGluGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGln 311  
 DB 1093 GGAGTCAAGAAAAGAAAGTGTGTGGGAGGAGCGGTGCG----- 1134  
 QY 312 AlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerSerAla 331  
 DB 1135 GGASTAGAGTAGAGGAGCAAAACAGAGAGATGAACATGTTCTGTGGAGTTCAGTGCT 1194  
 QY 332 SerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAla----- 348  
 DB 1195 TCTCCGGTGTGGAAAGCCACGCAAGAAATGCTATGACCAGAGGTCTCTCCACCGATGTA 1254  
 QY 349 -----AspHisAlaAspValLeuAlaLysGlyAlaGln 359  
 DB 1255 TCCACCGACCCCTAAAGTTTCTATTCTCTCTCACGCAACCTCGCTACTAAAGCGATGAG 1314  
 QY 360 AlaTyrAspGluTyr-----GlyArgAspAspTyrSerSerArgThrLysAsnGly 376  
 DB 1315 AATCTGTAGAGAACATGTCTACCGGAGAGAAAGGCGATGTGGAAATGGACCAACCGGT 1374  
 QY 377 SerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAla 396  
 DB 1375 ANTAACGGGGGGAAG-----TCACCTTACATGGGCAAAAAGGTAGGAC----- 1419  
 QY 397 GlnLeuTyrProLysAspGlyGluGlyArgAlaAlaValAlaMetProProAla 416  
 DB 1420 -----GTGGAAGACGGCGGTCCCGGTCTTAGGAAACAGCAGATGCCCGCGCG 1467  
 QY 417 SerValMetThrArgLeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsn 436  
 DB 1468 AGTGATGATGAGAGCAATAATCTGATTAATGTTGGAGAAACTCATTCGAAACCTTAAC 1527  
 QY 437 ThrTyrSerSerLeuIleGlyValValTrpSerLeuValSerTyrArgTTPGlyIleGlu 456  
 DB 1528 ACTTACTCTAGTCTTTGGCCCTTGTGGTCCCTTGTCTCTTCAAGTGGAAATATAAG 1587  
 QY 457 MetProAlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAla 476  
 DB 1588 ATGCCAAGATAATCAGTGGATCGATTTCGATATTAATCTGATGCTGTCTTGGATGCT 1647  
 QY 477 MetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLys 496

Db 1648 ATGTTTAGCTTGGTCTATTTATGCGCATTCACCAAGAAATATTGGTGGGAAATCA 1707  
 Qy 497 LeuAlaAlaIleAlaValMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAla 516  
 Db 1708 GTAGCAGGGTTGGCAGTCCGCTAGGTTCTTGACTGACACGCGGTGATCGCAGCCACC 1767  
 Qy 517 SerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeu 536  
 Db 1768 TCAATAGCAATGGTATTCGAGGTGATCTCTCCATATGCCATCGTTCAGGCTGCTCTT 1827  
 Qy 537 ProGlnGlyIleValProPheValPheAlaValGlyLeuValHisProAspIleLeu 556  
 Db 1828 CTTCAAGGATCGTTCTCTTTGTTTCCCAAGATATTAACGTCCATCTGATATCTC 1887  
 Qy 557 SerThrAla-----TyrGly-----ProIleThr 564  
 Db 1888 AGCACTGGCGTTATATTCGGAATGCTGGTCTTGTCTTGTCTGTAACA 1932  
 RESULT 15  
 AAZ57348  
 ID AAZ57348 standard; cDNA to mRNA; 2232 BP.  
 XX AAZ57348;  
 XX DT 03-APR-2000 (first entry)  
 XX DE Gravitropism stimulation response related protein AGR cDNA sequence.  
 XX KW Gravitropism stimulation response; AGR; plant root; ds.  
 XX OS Arabidopsis thaliana.  
 XX FH Location/Qualifiers  
 FT CDS  
 FT 24..1967  
 FT /tag= a  
 FT /product= "AGR"  
 FT /note= "a gravitropism stimulation response related  
 protein"  
 XX JP11318463-A.  
 XX FN 24-NOV-1999.  
 XX PD 15-MAY-1998; 98JP-00134097.  
 XX PF 15-MAY-1998; 98JP-00134097.  
 XX PR 15-MAY-1998; 98JP-00134097.  
 XX PA (KAZU-) ZH KAZUSA DNA KENKYUSHO.  
 XX PA (OJIP) OJI PAPER CO.  
 XX DR WPI: 2000-100768/09.  
 XX DR P-PSDB; AAY53129.  
 XX PT A gene participating to the gravitropism stimulation response of a plant  
 PT root.  
 XX PS Claim 2; Page 11-14; 20pp; Japanese.  
 XX CC The present sequence encodes a protein which participates in the  
 CC gravitropism stimulation response of a plant root. The protein is  
 CC designated AGR. AGR can be used to improve the fixing rate of a plant  
 CC root to soil  
 XX SQ Sequence 2232 BP; 601 A; 484 C; 544 G; 603 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.73e-98 Length: 2232  
 Score: 1535.00 Matches: 347  
 Percent Similarity: 60.89% Conservative: 64  
 Best Local Similarity: 51.41% Mismatches: 116  
 Query Match: 52.75% Indels: 148  
 Ds: 3 Gaps: 17

US-10-030-884-14 (1-573) x AAZ57348 (1-2232)  
 Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
 Db 24 ATGATCACCGCGCAAGACATGTACGATGTTTATAGCGGTATGGTGGCGTATACGTGCT 83  
 Qy 21 MetThrLeuAlaTyrGlySerValArgTTPArgIlePheThrProAspGlnCysSer 40  
 Db 84 ATGATATTAGCTATGTTCCGTACGTTGGTGGGATATTACACCGGACCAATGTTCC 143  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 144 GGTATAAACCGGTTCGTTGGGTTTTCCGGGTTTCCTCTCTCTCTCTCTCTCTCTCTCT 203  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIleVal 80  
 Db 204 TCCAAATGATCTTATGCAATGAATTACCACTTCTCTCGCTGCTGATCTCTTCAGAAAGTC 263  
 Qy 81 AlaValLeuAlaLeuLeu-----AlaLeuAlaSerArgGlyLeuSerSerPro 96  
 Db 264 GTTATCTCCCGCCTCTCTTTCTTGGCAGGCGTTTGGCGGCGAGGA----- 311  
 Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 312 -----AGCCTAGATGGATGATAGCGCTCTTTTCACTATCAACACTGCTTAACAGC 362  
 Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
 Db 363 TTGTAATGGGAATCCCATTTGCTTAGGGCGATGTCGGAGACTTCTCC---GGTAACCTA 419  
 Qy 137 MetValGlnValValValLeuGlnCysIleIleTyrThrLeuMetIlePheLeuPhe 156  
 Db 420 ATGTGTCAGATCGTGGTGTTCAGAGCATCATATGTTATGATTAATGCTCTCTCTTGT 479  
 Qy 157 GluTyrArgAlaAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
 Db 480 GAGTTCGTCGGGCTAAGCTTCTCATCTCCGAGAGTCTCCCGAG---ACGCTGTTCA 536  
 Qy 177 IleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGlu 195  
 Db 537 ATTACTTCTTCAGAGTTGACTCTGATGTTATCTCTTAATGGCGGTGAACCCCTCCAG 596  
 Qy 196 LeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAsp 215  
 Db 597 ACCGATCGGAGATA-----GGAGAC 617  
 Qy 216 AlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys 235  
 Db 618 GACGAAAGGTACACCGTGGTGTTCGAGATCAAGTGGCGCTCATCATGATCTCTTCA 677  
 Qy 236 -----SerHisSer-----HisSerGlnThrMetGlnProArgValSerAsn 249  
 Db 678 TTCACCAATCTCACGGCGAGGACTTAACTCTCTCATGATTAACCGCGAGCTTCAAT 737  
 Qy 250 LeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySer 269  
 Db 738 CTCACGGCGGTAGAGATTTACTCCGTTCAATCGTCACGAGAGCCGACGCGGAGAGCTTCT 797  
 Qy 270 SerPheAsnHisAlaAspPhePhe----- 277  
 Db 798 AGTTTAAATCAGACAGATTTCTACGCAATGTTTAACGCAAGCAAGCTCCAAGCCCTCGT 857  
 Qy 278 -----AsnIleValGlyAlaAlaLysGlyGly----- 288  
 Db 858 CACGGTTTACACTAATAGTACGCGCGCTGGAGTGGTCCAGTGGAGATGTTTACTCA 917  
 Qy 288 ----- 288  
 Db 918 CTTTCACTTCTTAAGCGCGTACGCGGAGAACGTCAAAATTTTGTATGAGAGTTATGAAG 977  
 Qy 288 ----- 288  
 Db 978 ACGCGCAAGAAAGCAGGAAGAGGAGGAGAGTATGATGGGGAAATTATACAACAATAAT 1037

QY 289 -----GlyAlaAla 291  
Db 1038 AGTGTTCGTCGTACCCACCGCCGAAACCCAAATGTTCAACGGGTCAACGAGTGGAGCAAGT 1097  
QY 292 GlyAspGluGlyGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGln 311  
Db 1098 GGAGTCAAGAAAGAAAGTGTGGCGGAGGAGCGGTGGC----- 1139  
QY 312 AlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAla 331  
Db 1140 GGAGTAGGAGTAGGAGGACAAACAGGAGATGAACATGTTCTGTTGGAGTTCGAGTGCT 1199  
QY 332 SerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAla 348  
Db 1200 TCTCCGGTGTGGGAAGCCACGCGAAGAACTATGACACGAGGTTCTTCCACCGATGTA 1259  
QY 349 -----AspHisAlaAspValLeuAlaLysGlyAlaGln 359  
Db 1260 TCCACCGACCTAAAGTTTCTATCTCTCACGACAACTCGCTACTAAAGCGATGCAG 1319  
QY 360 AlaTyrAspGluTyr-----GlyArgAspAspTyrSerSerArgThrLysAsnGly 376  
Db 1320 AATCTGATAGAACATGTCACCGGAGAAAGGCGATGTGGAAATGGACCAAGCGGT 1379  
QY 377 SerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAla 396  
Db 1380 AATAACGGGGGAAG-----TCACCTTACATGGGCAAAAAGGTAGCGAC----- 1424  
QY 397 GlnLeuTyrProLysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProAla 416  
Db 1425 -----GTGGAAGACGGCGGTCCCGTCTAGGAAACAGCAGATGCCCGCGCG 1472  
QY 417 SerValMetThrArgLeuLeuLeuMetValTrpArgLysLeuLeuArgAsnProAsn 436  
Db 1473 AGTGTGATGACGAGCACTAATCTCATATGTTTGGAGAAACTCATTCGAAACCCCTAAC 1532  
QY 437 ThrTyrSerSerLeuLeuGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGlu 456  
Db 1533 ACTTACTCTAGTCTTTGGGCTTGTGTCCTTGTCTCTTCAAGTGGAAATATAAG 1592  
QY 457 MetProAlaIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAla 476  
Db 1593 ATGCCACGATAANTGAGTGGATCGATTTTCGATATTTATCTGATGCTGGTCTTGGATGGCT 1652  
QY 477 MetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLys 496  
Db 1653 ARGTTTGTCTTGGTCTATTATGGCATTCGACCAAGATTAATGCGTGGGAAATCA 1712  
QY 497 LeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAla 516  
Db 1713 GTAGCAGGTTTGGCATGCCGTAAAGTTCTTGTGACCGACCGCGTATCGCAGCCACC 1772  
QY 517 SerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeu 536  
Db 1773 TCAATAGCAATTTGGTATTTCGAGGTGATCTCTCCATATCGGCATCGTTCAGGCTGCTTT 1832  
QY 537 ProGlnGlyIleValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeu 556  
Db 1833 CTTCAAGGAATGTTCTTCTTTTTCGCAAGAAATATAAGTCCATCCATGATATCTC 1892  
QY 557 SerThrAla-----TyrGly-----ProIleThr 564  
Db 1893 AGCACTGCGGTATATTCGAATGTGGTGTGTTTGGCTGTGAACA 1937

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 3, 2004, 09:09:54 ; Search time 6280 Seconds

(without alignments)  
3954.707 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910

Sequence: 1 MITALDIYLVITAVVPLVA.....DILSTAYGPITSHGFITCHS 573

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.epool/US10030884/runat\_24022004\_103922\_1948/app\_query.fasta\_1.711  
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.19e -MINMATCH=0.1 -LOOPCL=0 -LCORPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdl -LIST=45  
-DOCALIGN=200 -THR SCORE=spect -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10030884@cgn 1 1 2586 @runat\_24022004\_103922\_1948 -NCPU=3  
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vl.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.fod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	ID	Description
1	2910	100.0	2769	6	AX046826 Sequence
2	2137.5	73.5	2233	8	AK103181 Oryza sat
3	2129.5	73.2	1926	6	AX653844 Sequence
4	2020.5	69.4	1686	6	AX652923 Sequence
5	1977	67.9	91883	8	CNS080CDY
6	1959.5	67.3	141025	2	AC123528 Oryza sat
7	1785.5	61.4	2297	8	AF515435 Populus t
8	1685	57.2	2358	8	AF056027 Oryza sat
9	1665	57.2	2402	8	AK102343 Oryza sat
10	1659.5	57.0	2457	8	AK101504 Oryza sat
11	1656	56.9	2343	8	AK099634 Oryza sat
12	1655.5	56.9	2126	8	AK063976 Oryza sat
13	1655.5	56.9	2470	8	AK103208 Oryza sat
14	1647.5	56.6	1827	6	AX654121 Sequence
15	1647.5	56.6	1869	8	AY093960 Arabidops
16	1647.5	56.6	2262	8	AF372950 Arabidops
17	1647	56.6	2262	8	AB085897 Cucumis s
18	1643	56.5	2549	6	AX046842 Sequence
19	1641.5	56.4	2276	8	AF089084 Arabidops
20	1641.5	56.4	2292	6	A68597 Sequence 11
21	1641.5	56.4	2292	6	A79382 Sequence 11
22	1623.5	55.8	9561	2	AC135795 Medicago
23	1614.5	55.5	2473	8	AF190881 Populus t
24	1605.5	55.2	2790	8	AF515434 Populus t
25	1599	54.9	1952	8	AY302060 Populus t
26	1596	54.8	1845	6	AX652849 Sequence
27	1592	54.7	1884	6	AX653686 Sequence
28	1587	54.5	1800	8	AY222857 Pisum sat
29	1571.5	54.0	2276	8	AK101191 Oryza sat
30	1569	53.9	2522	8	AB112364 Pisum sat
31	1566.5	53.8	2293	6	AX046850 Sequence
32	1566.5	53.8	2293	8	BT008949 Trifolium
33	1560.5	53.6	1923	6	AX506152 Sequence
34	1560.5	53.6	2380	8	BT002085 Arabidops
35	1560.5	53.6	2489	8	AY136327 Arabidops
36	1560.5	53.6	2499	8	AF087818 Arabidops
37	1558.5	53.6	2462	8	AJ249237 Brassica
38	1542.5	53.0	2358	8	BT008691 Arabidops
39	1542.5	53.0	2404	8	AY091009 Arabidops
40	1541.5	53.0	2364	8	AF087016 Arabidops
41	1539.5	52.9	91870	8	AP004800 Oryza sat
42	1539	52.9	2316	8	BTJ249298 Brassica
43	1535	52.7	1944	8	BT001152 Arabidops
44	1535	52.7	2196	8	AF056026 Arabidops
45	1535	52.7	2232	6	E31424 Gene partic

ALIGNMENTS

RESULT 1

AX046826  
 LOCUS AX046826 2769 bp DNA linear PAT 15-DEC-2000  
 DEFINITION Sequence 13 from Patent WO0068389.  
 ACCESSION AX046826  
 VERSION AX046826.1 GI:11876320  
 KEYWORDS Zea mays  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 Clade; Panicoideae; Andropogoneae; Zea.  
 1  
 ORZOCO, B.M., MENG, Z., BRUCE, W.B., CAHOON, R.E. and TAO, Y.  
 Auxin transport proteins  
 Patent: WO 0068389-A 13 16-NOV-2000;  
 E.I. DU PONT DE NEMOURS AND COMPANY (US); Pioneer Hi-Bred  
 International, Inc. (US)  
 FEATURES  
 source  
 Location/Qualifiers  
 1..2769  
 /organism="Zea mays"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4577"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,5e-155 Length: 2769  
 Score: 2910.00 Matches: 573  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-030-884-14 (1-573) x AX046826 (1-2769)  
 QY 1 MetIleThAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
 Db 171 ATGATCACCGCGTGGACCTCTACCACTGCTGACCGCGTGTGCGGTGTACGTGGCC 230  
 QY 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
 Db 231 ATGACGTGGCGTACGGCTCCGTCCTGTGGCGCATCTTACGCGGACCACTGCTCC 290  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 291 GGGATCAACCGCTTCGTGGCGCTTTCGCGTGGCGCTCCTCTCTCTTCCATCTCATCC 350  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIleVal 80  
 Db 351 ACCAAGACCCCTTCGCCATGAACCTTCGCGCTTCTCGCGCGCGACACGCTGCAGAGGTG 410  
 QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 Db 411 GCGTCTCTCGCGTGTGGCGCTGGCGCTCCGCGGCTCTCTCTCCGCGCGCTCGGG 470  
 QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
 Db 471 CTCGACTGGAGCATCAGCTCTTCTCCCTCCACGCTCCCAACACGCTCTGTCATGGGC 530  
 QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 Db 531 ATCCCGTCTGCGAGCATGTACGGCGCTGCTCGCGCGGACGCTCATGTGTCCAGTTC 590  
 QY 141 ValValLeuGlnCysIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 160  
 Db 591 GTGCTCTCCAGTGATCATCTGGTACACGCTATGCTCTTCTCTCTCTCTCTCTCTCTCT 650  
 QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 Db 651 GCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 710  
 QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
 Db 711 CGGCTGACTCCGACGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770

QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
 Db 771 GACGGGTGCGCGCGCGCGCGCTCTCTCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 830  
 QY 221 ValThrValArgIleSerThrSerArgSerGluAlaAlaCysSerHisSerHisSer 240  
 Db 831 GTACCGGTGGCAAGTCCACAGCTCGCTCGAGCGCGCGCTCGCTCGCTCGCTCGCTCG 890  
 QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260  
 Db 891 CAGACCATGCGAGCCCGGTGTCCAACTCTCCGCGGTGGAGATCTACTCGCTGCAGTCG 950  
 QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleVal 280  
 Db 951 TCGCGCAACCCACCCCGCGCGGTCCAGCTTCAACACCCGCTCTTCAACATCGTC 1010  
 QY 281 GlyAlaAlaAlaIleGlyGlyGlyGlyGlyAlaAlaGlyAspGluGluIleGlyAlaCysGly 300  
 Db 1011 GCG 1070  
 QY 301 GlyGlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAlaIleArgLys 320  
 Db 1071 GCGCGCGCGCGAGACACTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1130  
 QY 321 AspLeuHisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaAlaVal 340  
 Db 1131 GACCTGCACATGCTCTGCTGAGCTCCAGCGCTCGCGCGCTCGCGCGCGCGCGCGCG 1190  
 QY 341 HisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaIleGlyAlaGlnAla 360  
 Db 1191 CACGTCTTCCGCGCGCGCGCGCTGACCATGCGCGCGCTCTCCCAAGAGCGCGCGCG 1250  
 QY 361 TyrAspGluTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyVala 380  
 Db 1251 TACACGAGTACCGCGCGCGCGCGCTAGCTACAGCAGCAGCAGCAGCAGCAGCAGCAG 1310  
 QY 381 AspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrPro 400  
 Db 1311 GACAAGCGCGCGCGCGCGCTGTGCAAGCTGGGCTTCAACTCGACGCGCGCGCTGTATCCC 1370  
 QY 401 LysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThr 420  
 Db 1371 AAGACACGCGCGAGGCGCGCGCGCGCTGCGCATGCGCGCGCGCGCGCGCGCGCGCG 1430  
 QY 421 ArgLeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSer 440  
 Db 1431 CGGCTCATCTCATCTGATGTGTGAGGAAGCTGATCCGGAACCCCAACACCTACTCCAGC 1490  
 QY 441 LeuIleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIle 460  
 Db 1491 CTCATCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1550  
 QY 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480  
 Db 1551 ATCCCGCGTGGATTCGATCTCTGATCGCGCGGCTCTCGGATGGCCATGTTTCAGCCTA 1610  
 QY 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaAlaIle 500  
 Db 1611 GGCCTGTTTCGCGCTGCGAGCGCGAGGATCATCGCGTGGGGAACAAGCTGGCGCGCATC 1670  
 QY 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaVal 520  
 Db 1671 CGGATGGCGCTCGGTGCTGCGAGCGCGCGCGGTCTGCGCGCGCGCTCCATCGCGCTC 1730  
 QY 521 GlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaLeuProGlnGlyIle 540  
 Db 1731 GGTCTGCGCGCGCTCTCTCCATCGCATCTGCTCCAGGCTGCTGCTGCTGCTGCTGCTG 1790  
 QY 541 ValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAlaTyr 560  
 Db 1791 GTGCGCTTGTGTTTCCGCGAGAGTACGGCTTATCCCGACATCTTCCGACACGCGCTAT 1850

Qy 561 GlyProfileThrSerHisGlyPheIleThrCysHisSer 573  
 Db 1851 GGTCCATAACATCGCATGGTTTCATCATCTGCCATAGT 1889

RESULT 2  
 AK103181 2233 bp mRNA linear PLN 24-JUL-2003  
 LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J033121124, full  
 DEFINITION insert sequence.  
 ACCESSION AK103181  
 VERSION AK103181.1 GI:32988390  
 KEYWORDS FLI cDNA; CAP trapper.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

REFERENCE  
 1  
 AUTHORS The Rice Full-Length cDNA Consortium, National Institute of  
 Agrobiological Sciences Rice Full-length cDNA Project Team,  
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
 Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K.,  
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
 Kurotsaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
 Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,  
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;  
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y.,  
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from  
 japonica rice  
 JOURNAL Science 301 (5631), 376-379 (2003)  
 MEDLINE 22752273  
 PUBMED 12869764

REFERENCE  
 2 (bases 1 to 2233)  
 AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoaka, T.,  
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Kurotsaki, T., Kusumegi, T., Li, C., Lu, M.,  
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
 Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,  
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
 Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,  
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and  
 Yoshimura, A.  
 Direct submission  
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of  
 Agrobiological Sciences, Department of Molecular Genetics, Head of  
 Laboratory of Gene Expression; 2-1-2 Kannondai, Teukuba, Ibaraki  
 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp,  
 Tel:81-29-838-7007, Fax:81-29-838-7007]  
 This clone is one of the 28K full-length cDNA clones from japonica  
 rice.  
 URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,  
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,

Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and  
 Yamamoto, M.  
 FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,  
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
 Kodama, T., Kurotsaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
 Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,  
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
 Yoshimura, A., Matsubara, K. and Murakami, K.  
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
 Hiraoaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kouda, M.,  
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,  
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
 Yasunishi, A. and Hayashizaki, Y.  
 Location/Qualifiers  
 1. 2233  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J033121124"

ORIGIN  
 Alignment Scores: Length: 2233  
 Pred. No.: 1,19e-111 Matches: 457  
 Score: 2137.50 Conservat: 26  
 Percent Similarity: 81.86% Mismatches: 36  
 Best Local Similarity: 77.46% Indels: 71  
 Query Match: 73.45% Gaps: 17  
 DB: 8

US-10-030-884-14 (1-573) x AK103181 (1-2233)  
 QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
 Db 222 ATGATCACCAGTGGTGACCTTACACACGCTTCACGCGGCTGGTGCCTGTACGTGGG 281  
 QY 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgPheThrProAspGlnCysSer 40  
 Db 282 ATGACGCTGGGTACGCTCCGTCGGTGGTGGCGCATCTTCTCCCGGACGAGTCTCC 341  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 342 GGCATCAACCGCTTCGTCGGCTTTCGCGGCTCTCTCTCTCCCTTCACCTTCATCTCC 401  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIysVal 80  
 Db 402 ACCACACACCCCTTCGCCATGAACCTTCCTCCGCGCGCGACACGCTCCAGAAGCTC 461  
 QY 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
 Db 462 ATCGTCTCGCCCTCTCTCGCCCTCTGTGTCGCGCTCTCCGCGCGCGGG----- 509  
 QY 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 510 -----TCATCGACTGGCTGCATCACCTCTTCTCCCTCTCCACCTCCCAACACC 560  
 QY 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSer-----Ala 133  
 Db 561 CTCGTATGGGATCCCGCTGTCTCAAGGGGATGTACGCGCGCGCGCGCGCGCTGACTCC 620  
 QY 134 GlyThrLeuMetValGlnValValLeuGlnCysIleIleIleTrpTyrThrLeuMetLeu 153  
 Db 621 GGCAGGCTCATGTGTGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 680

154 PheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAla 173  
Db TTCTCTTCGAGATACCGCGCGCGCGCTTCTGTGTCATGAGCAGATTCCCGAC---ACC 737  
174 AlaAlaSerIleValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAsp 193  
Db GCCGCTCCATCGTCTCTTCGCGGTGGATTCGACGTCGTCTCGCTCGCCGCGCGCGT 797  
194 -----ValGluLeuGluAlaGluProAspGlyValAlaAlaGlyAlaValSer 210  
Db GTGTGAGCCGCGAGCTCGAGCGGAG----- 824  
211 SerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThrSerSerArg 230  
Db GCGGAGGTGCGGAGCATGGCAGGATCGCGGTCACTGTCGCAAGTCGACGAGCTCGGCG 894  
231 SerGluAlaAlaCysSerHis-----SerHisSerGlnThrMetGlnProArgVal 247  
Db TCCGAGGCGCGGTGCTCGACGGGACGAGTCACTCGCAGTCTCATGCGCGCGCGT 944  
248 SerAsnLeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArg 267  
Db TCCAAACCTCTCGCGGTGCGAGATTACTCTGCTCAGCTGCTCGCGAACCACGACCGCGT 1004  
268 GlySerSerPheAsnHisAlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyVal 287  
Db GGCTCCAGCTTCAACACCGCGGAGTCTTCAACATCGTC-----GGC 1046  
288 GlyGlyAlaAlaGlyAspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyHisSer 307  
Db ACGGCACGACGCGGCGACGAGGAGAGGGCGCC-----GCCGGTGGCGGGCGCACTCG 1100  
308 ProGlnProGlnAlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrp 327  
Db CGCGACGCG-----GTGGTGGGAGAGGAGGACCTGCACATGTTGCTGTGG 1148  
328 SerSerSerAlaSerProValSerGluArgAla-----AlaValHis 341  
Db ASCTCAAGCGCTCGCGGTGTGAGACGCGCGCGCAGCAGCAGTGGCGGTGCAC 1208  
342 ValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTyr 361  
Db GTCTTCGCGGTGGTGGCGCGCACACGCGCAC-----GCCAAAGAGAGCTCAGCGCTAT 1262  
362 AspGluTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySerGlyAlaAsp 381  
Db GATGAG-----TACAGTTCCGGAAACAAGAC-----GAG 1292  
382 LysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLys 401  
Db AAGACGCGGCGCGCGCTGTCCAGCTGGGTGCCAATCGACGCGCGCAGCTCGGCGCAAG 1352  
402 AspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThrArg 421  
Db GACGACGCGGAGGGGATGCGCGG-----GCCATCCCGCGCGAGCTGATGACGAGG 1406  
422 LeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeu 441  
Db CTCATCTCGATCATGGTTTGAGAGAGCTGATCAGAACCCCAACACTTACTCCAGCGCTC 1466  
442 IleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIleIle 461  
Db CTGTGTGTCATCTGTCCTCGTCTCTACAGGTGGGGATTGAGATGCCCATCATC 1526  
462 AlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGly 481  
Db GCCCGGTGATTTGATCTCTGTCAGTGCAGGCTTGGAAATGCCATGTTTCAGCTTAGGA 1586  
482 LeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAla 501  
Db TTGTTTCATGGATTGCAGCCACGATCATTCCTGTGGGAATCCCTGTCTGCTGATGCC 1546

QY	502	MetGlyValAlaArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGly	521
DB	1647	ATGGCCGTCCGGTCTCTTCGCGGGTCCGGCGGTCTATGCTCCGCCCTCCATCGCGCGTGG	1706
QY	522	LeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaLeuProGlnGlyIleVal	541
DB	1707	CTTCGCGGGTGTCTTCGACATGTCATGTTTCAGGCGCTCTTCTCAAGGAATCGTG	1766
QY	542	ProPheValPheAlaIysGluTyrGlyValHisProAspIleLeuSerThrAla	559
DB	1767	CCCTTTGTGTTCGAAGGAGTAGTAATGTTTCATCTTAAACATTCTGACACAGCGTGATA	1826
QY	560	TyrGly	564
DB	1827	TTCGGGATGCTGATGCTCTCTCCCATCACA	1856
RESULT 3			
AX653844			
LOCUS	AX653844	1926 bp	DNA
DEFINITION	Sequence 3714 from Patent WO03000898.		
ACCESSION	AX653844		
VERSION	AX653844.1	GI:29156658	
KEYWORDS			
SOURCE	Oryza sativa		
ORGANISM	Oryza sativa		
REFERENCE			
AUTHORS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzoae; Oryza.		
TITLE	Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M., Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.		
JOURNAL	Plant Genes Involved in defense against pathogens		
SYNGENTA PARTICIPATIONS AG (CH)	Patent: WO 03000898-A 3714 03-JAN-2003;		
FEATURES			
source	1. 1926		
ORIGIN	/organism="Oryza sativa"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:4530"		
Alignment Scores:			
Pred. No.:	2,81e-111	Length:	1926
Score:	2129.50	Matches:	456
Percent Similarity:	77.58%	Conservative:	23
Best Local Similarity:	73.87%	Mismatches:	42
Query Match:	73.18%	Indels:	97
DB:	6	Gaps:	16
US-10-030-884-14 (1-573) x AX653844 (1-1926)			
QY	1	MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla	20
DB	1	ATGATCACGGTGTGGACCTGTACCAAGTCTGACGGCGGTGGTGCCTGTACGTGGCG	60
QY	21	MetThrLeuAlaTyrGlySerValArgTyrTpArgIlePheThrProAspGlnCysSer	40
DB	61	ATCACGCTGGGACGACCTCCGTCGGTGGTGCGATCTTCTCCCGGACCCAGTGTCTC	120
QY	41	GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheLeuSer	60
DB	121	GGCATCAACGGCTTCGTCCGCTCTTCGGCGTCCCGCTCTCTCTCCACTTCATCTCC	180
QY	61	ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIysVal	80
DB	181	ACCAACAACCCCTTCGCCATGAACCTCCGCTTCTCCCGCGCGACACGCTCCAGAAGCTC	240
QY	81	AlaValLeuAlaLeuLeuAla	96
DB	241	ATCGTCTTCGCCCTCTCGCSTCTGGTGGCGCGCTCTCCGCGCGGG	288
QY	97	ArgAlaLeuGlyLeuAspTyrPserIleThrLeuPheSerLeuSerThrLeuProAsnThr	116

Db 289 -----TCMCTCGACTGGCTCATCACCCCTCTTCTCCCTCTCCACCCCTCCCAACACC 339  
 Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAla----- 133  
 Db 340 CTCGTATGGCATCCCGCTCTCAAGGGATGTACCGCCGCGCGCGCGCGCGCGCGCG 399  
 Qy 134 -----GlyThrLeuMetValIleValValValLeuGlnCysIleIleTyrThr 150  
 Db 400 GCCGACTCCGCGCGCTCATCGTGCAGATCGTCTGCTCCAGTGCATCATCTGGTACAG 459  
 Qy 151 LeuMetLeuPheLeuPheGlyTyrArgAlaAlaAlaAlaValLeuValLeuAspGlnPhePro 170  
 Db 460 CTCATGCTGTCTCTTCAGTACCGYGGCGCGCGGCTTCTTGTATGATGAGCAGATTCGCC 519  
 Qy 171 AspGlyAlaAlaAlaSerIleValSerPheArgValAspSerValValSerLeuAla 190  
 Db 520 GAC---ACCGCGCGCTCATCGTCTCTTCCTCGGGTGGATTCCGACGTCGCTCGCTGCC 576  
 Qy 191 ArgGlyAsp-----ValGluLeuGluAlaGluProAspGlyValAlaGlyAlaGly 207  
 Db 577 GCGCGCGTGTGGAGCCGCGGAGCTGCGAGCGGAG----- 612  
 Qy 208 AlaValSerSerArgGlyGlyAspAlaGlyArgValArgValThrValArgLyseThr 227  
 Db 613 -----GCGGAGTTCGGGCGATGCGAGGATGCGGGTCAACCGTGCAGAGTCGACG 663  
 Qy 228 SerSerArgSerGluAlaAlaCysSerHis-----SerHisSerGlnThrMetGln 244  
 Db 664 AGTCCGCGCTCGAGCGCGGCTGCTCGACCGGAGCGCAGTCACACTCGCAGTCCATGCG 723  
 Qy 245 ProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnPro 264  
 Db 724 CCGCGCTCTCAACCTCTCGCGCTCGAGATTTACTCGTGCAGTGTGTCGCGAAGCCCG 783  
 Qy 265 ThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleValGlyAlaAlaAla 284  
 Db 784 AGCGCGGTGGCTCAGCTTCAACCGCCGAGTCTTCAACATCGTC----- 831  
 Qy 285 LysGlyGlyGlyAlaAlaGlyAspGluLysGlyAlaCysGlyGlyGlyGlyGly 304  
 Db 832 -----GGCAACGGCAAGCAGCGGCGACGAGGAGGAGGCGCC-----GCGGTGGCGGG 879  
 Qy 305 GlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLysAspLeuHisMet 324  
 Db 880 GGCCACTCGCCGACCGC-----GTGGTGGGAGAGAGGAGGACCTGCACATG 927  
 Qy 325 LeuValTrpSerSerAlaSerProValSerGluArgAla----- 338  
 Db 928 TTCGTGTGAGCTCAAGCGCTCGCGGTGTCGAGCGCGCGCGCAGCAGCTGGC 987  
 Qy 339 AlaValHisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaLysGlyAla 358  
 Db 988 GCGGTGCAGCTTTCGCGGTGTGTGCGCGGACCGACCGCGAC-----GCCAAGGAGCT 1041  
 Qy 359 GlnAlaLysAspGluTyrGlyArgAspTyrSerSerArgThrLysAsnGlySerGly 378  
 Db 1042 CAGGCTATGATGAG-----TACAGTTTCGGGAACAAGAC----- 1077  
 Qy 379 GlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeu 398  
 Db 1078 -----GAGAGGACGGCGCGAGCTGTCCAAGCTGGGTCCAACTCGACGCGCGAGCTC 1131  
 Qy 399 TyrProLysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerVal 418  
 Db 1132 CGGCCAAGAGGACGCGCGGAGGATGCGCGG-----GCATCGCCCGCGCGAGCTG 1185  
 Qy 419 MetThrArgLeuIleLeuMetValTrpArgLysLeuIleArgAsnProAsnThrTyr 438  
 Db 1186 ATGACGAGCTCATCTGATCATGTGTTTGGAGAGCTGATCAGGAACCCCAACACTTAC 1245  
 Qy 439 SerSerLeuGlyValValTrpSerLeuValSerTyrArgTyrGlyIleGluMetPro 458  
 Db 1246 TCCAGGCTCTCTGGTGTCTATCTGGTCCCTCTCTCTACAGGTGGGGATTGAGATGCA 1305

Qy 459 AlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPhe 478  
 Db 1306 GCATCATCGCCCGGTGCAATTCGATCTGTGATGCAGGCTTGGAAATGSCCATGTTC 1365  
 Qy 479 SerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAla 498  
 Db 1366 AGCTAGGATTTTCATGGCATTCGAGCCACCGATCATTCGCTGGGAACCTCCCTTGT 1425  
 Qy 499 AlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerIle 518  
 Db 1426 TGTATGCCATGGCGTCAAGTTCCTGCGGTCCGCGCTCATGGCTCGCGCTCCATC 1485  
 Qy 519 AlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleVal----- 532  
 Db 1486 GCGGTGCGGATTCGCGGGTGTCTTGCACATTCGCAATTCGATTCAGTTCAGTTCAGTTC 1545  
 Qy 532 ----- 532  
 Db 1546 AGCTTCTGCACAAATGCAATCTTACTGCAAAAGATCTATCTGTAACCTCAACATTCG 1605  
 Qy 533 -----GlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlu 548  
 Db 1606 TCATTGAATTCCTCAGCGCGCTCTTCTCAAGGAATCGTGCCTTGTGTTGCCAAGGAG 1665  
 Qy 549 TyrGlyValHisProAspIleLeuSerThrAlaTyrGlyProIleThrSerHisGlyPhe 568  
 Db 1666 TACAATGTTTCATCTTAACATCTGAGCACGCGTATGATCTCTTACCATCTTTCTTT 1725  
 RESULT 4  
 AX652923 1686 bp DNA linear PAT 22-MAR-2003  
 LOCUS AX652923  
 DEFINITION Sequence 2793 from Patent WO03000898.  
 ACCESSION AX652923  
 VERSION AX652923.1 GI:29155737  
 KEYWORDS  
 SOURCE Oryza sativa  
 ORGANISM Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE 1  
 AUTHORS Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M.,  
 Katagiri, F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
 TITLE Plant genes involved in defense against pathogens  
 JOURNAL Patent: WO 03000898-A 2793 03-JAN-2003;  
 SYNGENTA PARTICIPATIONS AG (CH)  
 FEATURES  
 source  
 1. 1686  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4530"  
 ORIGIN  
 Alignment Scores: 3.32e-105 Length: 1686  
 Pred. No.: 2020.50 Matches: 442  
 Score: 2020.50  
 Percent Similarity: 75.69% Conservative: 22  
 Best Local Similarity: 72.10% Mismatches: 39  
 Query Match: 69.43% Indels: 111  
 DB: 6 Gaps: 16  
 US-10-030-884-14 (1-573) x AX652923 (1-1686)  
 Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 Db 1 ATGATCAGCGTGTGGAGCCCTGATCCAGCTCTGACGCGGTGTGCGGTGTGACGTGGCG 60  
 Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
 Db 61 ATGACGCTGGGTACGCTCCGTCGCGGTGGTGGCGCATCTCTCCCGGACCATGTCCTCC 120  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60

Db 121 GGCATCAACCGCTTGGTGGCGCTCTTCCGCGTCCGCTCTCTCTCTCCATCTTCATCTCC 180  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAspThrLeuGlnLysVal 80  
Db 181 ACCAACACCGCTTCCGCTCAACCTCCGCTTCTCTCCGCGCGGACACGCTCCAGAGCTC 240  
Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
Db 241 ATCGTCTCCGCGCTCTCTCCGCTCTGGTGGCGCTCTCCGCGCGGG----- 288  
Qy 97 ArgAlaLeuGlyLeuAspTrpSerPheThrLeuPheSerLeuSer-ThrLeuProAsnTh 116  
Db 289 -----TCCCTCGACTGGCTCATCACCTCTCTCTCCCTCTCAACCTTCCCAAMAC 339  
Qy 116 rleuValMetGlyLysProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLe 136  
Db 340 CCTCGTCATGGCAGTCCCGCTCTCAAGGAT----- 372  
Qy 136 uMetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPh 156  
Db 373 -----TGCATCATCTGGTACACGCTCATGCTGTTCTCTT 407  
Qy 156 eGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSe 176  
Db 408 CGAGTACCGTGGCGCGCGCTTCTTGTATGAGGAGTCCCGGAC-----ACGCGCGCTC 464  
Qy 176 rIleValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAsp----- 193  
Db 465 CATCGTCTCTTCCGGTGGATCCGAGTGTCTCGCTCGCGCGCGCGTGGTGGAGC 524  
Qy 194 -ValGlnLeuGluAlaGluProAspGlyValAlaGlyAlaValSerSerArgG 213  
Db 525 CGCGAGCTGCGAGCGGAG-----GCGAGGT 551  
Qy 213 yGlyAspAlaGlyArgValArgValThrValArgLysSerThrSerArgSerGluAl 233  
Db 552 CGGGGACGATGCAAGATCGAGTCCAGTCCGTCGCGAAGTCCGAGAGCTCGGCTCCGAGG 611  
Qy 233 alaCysSerHis-----SerHisSerGlnThrMetGlnProArgValSerAsnLe 250  
Db 612 GCGGTCTCGCACGGGACGAGTCAACATCGAGTCCATGCGCGCGCGCTCCCAAGCT 671  
Qy 250 uSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSe 270  
Db 672 CTCGGCGTGCAGATTACTCTGCTGAGTCTGTCGGAACCGCGAGCGCGGGCTCCAG 731  
Qy 270 rPheAsnHisAlaAspPheAsnIleValGlyAlaAlaLysGlyGlyGlyAl 290  
Db 732 CTTCAACACCGCGAGTCTTCAACATCGTC-----GGCAACGGCAA 773  
Qy 290 aAlaGlyAspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnPr 310  
Db 774 GCACGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 827  
Qy 310 oGlnAlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerSe 330  
Db 828 G-----GTGTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 875  
Qy 330 rAlaSerProValSerGluArgAla-----AlaValHisValPh 343  
Db 876 CGCTCGCGCGTTCGAGCT 935  
Qy 343 eGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTrpAspG 363  
Db 936 CGCGCGTGGTGGCGCGGACGACGCGGAC-----GCCAAGGTGCTCAGGCTTATGATGA 989  
Qy 363 uTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysG 383  
Db 990 G-----TACAGTTCGGGAGCAAGAA-----GAGAGGA 1019  
Qy 383 yGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAs 403

Db 1020 CGGCGCGCACACTGTCCAAAGCTGGGTCCAACTCGACGGCGAGCTCCGGCCAAAGACGA 1079  
Qy 403 pGlyGluGlyArgAlaAlaAlaValAlaMetProAlaSerValMetThrArgLeuI 423  
Db 1080 CGGCGAGGGAGGGGGCA-----GCGATGCCGCGCGGAGCGTGTATGACGAGGCTCAT 1133  
Qy 423 eLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleG 443  
Db 1134 ACTGATCATGGTCTGGAGGAGCTGATCAGGAACCCAAACACTTACTCCAGGCTCTCTCG 1193  
Qy 443 yValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetProAlaIleAlaAla 463  
Db 1194 TGTCTATCTGCTCTCTCTCTCTCTACAGGTGGGAATTGAGATGCCGCTATCATGCCCG 1253  
Qy 463 gSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPh 483  
Db 1254 GTCGATTCGATCTCTTTCAGATGCGAGGCTCGGAATGGCCATGTTTCAGCTAGGATGTT 1313  
Qy 483 eMetAlaLeuGlnProArgIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetG 503  
Db 1314 CATGGCATTTGCACCCAGGATCATTCCTGTGGGAACCTCCCTTGCCTTCGTATGCCATGGC 1373  
Qy 503 yValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuAr 523  
Db 1374 CGTCAGGTTCTCTGTTGGTCTCTGCTCATGGCTGGGCTCCATCGCGCTCCATCGCGTGG 1433  
Qy 523 gGlyValLeuLeuHisIleAlaIleVal----- 532  
Db 1434 CGGGGTCTCTTCACATTCGCAATTCGAGTAACTGCAAGTTCAGGCTTCGACCA 1493  
Qy 533 -----G 533  
Db 1494 AAATGTCATCTTAACTCGCAAAAGATCTATCTGTAACCTCAACATTCGATTCATTC 1553  
Qy 533 nAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlyTyrGlyValHisPr 553  
Db 1554 GCGCGCTCTTCTTCCCAAGGAATCGTCCCTTCGTTGTTGTCAGGAGTACATGTTTCATCC 1613  
Qy 553 oAspIleLeuSerThrAlaTyrGlyProIleThrSer 565  
Db 1614 TAACATTCGAGCACAGCGTATGATCTCTTTACCAC 1650  
RESULT 5  
CNS08CDY 91883 bp DNA linear PLN 21-NOV-2003  
LOCUS Oryza sativa chromosome 12, . BAC OJ1575.G05 of library Monsanto  
DEFINITION from chromosome 12 of cultivar Nipponbare of ssp. japonica of Oryza  
sativa (rice), complete sequence.  
ACCESSION BX000508.1 GI:25900581  
VERSION BX000508.1  
KEYWORDS HTG.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriatoideae; Oryzaceae; Oryza.  
REFERENCE 1 (bases 1 to 91883)  
Choisne, N., Orjeda, G., Cattolico, L., Demange, N., Mincker, P.,  
Segurens, B., Pelletier, E., Scarpelli, C., Salanoubat, M.,  
Weissenbach, J. and Quetier, F.  
Oryza sativa chromosome 12 sequencing  
unpublished  
2 (bases 1 to 91883)  
Genoscope.  
Direct Submission  
Submitted (20-NOV-2003) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
Center code: GS  
Center code: GS  
Web site: http://www.genoscope.cns.fr/  
Contact: Seqref@genoscope.cns.fr

```

The following sequence is oriented from the T7 to the SP6 end. The
nucleotide sequence of this BAC clone was generated by combining
Monsanto and Genoscope sequencing data.
Upsream BAC (overlapping the T7 end) : OSUNB8A009P13 (AC=BX000498)
Downstream BAC (overlapping the SP6 end) : QJ1239_F07 (AC=BX000504)
----- Finishing boundaries -----
FINISHED SEGMENT STARTS AT BASE 28300
FINISHED SEGMENT ENDS AT BASE 69200

```

## FEATURES

```
1. .91883
/organism="Oryza sativa"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="12"
/clone="OJ1575_G05"
/clone_lib="Monsanto"
```

## ORIGIN

Alignment Scores:

1.23e-100	Length:	91883
Pred. No.:	Matches:	458
Score:	1977.00	
Percent Similarity:	59.34%	Conservative: 25
Best Local Similarity:	56.27%	Mismatches: 39
Query Match:	67.94%	Indels: 292
DB:	8	Gaps: 19

US-10-030-884-14 (1-573) x CNS08CDY (1-91883)

Qy	1	MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla	20
Db	32692	ATGATCACGGTGGGAGCCTGTACCACTGCTGACGGGGTGGTGGCTGTGTACGTGGCG	32751
Qy	21	MetThrLeuAlaTyrGlySerValArgTTPtApArgIlePheThrProAspGlnCysSer	40
Db	32752	ATGACGCTGGGCTACGCTCCGCTCCGGTGGTGGCGCATCTTCTCCCGGACCGAGTGCTCC	32811
Qy	41	GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer	60
Db	32812	GGCATCAACCGCTTCGCGGTCTTCGCGCGTCCGCGCTCTCTCTCTCCATTCATCTCC	32871
Qy	61	ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal	80
Db	32872	ACCAACACCCCTTCGCGCATGAACCTTCGCGTCTTCGCGCGGACACAGCTCCAGAGCTC	32931
Qy	81	AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro	96
Db	32932	ATCGTCTCGCCCTCTCTCGCGCTCTGGTCCGCGCTCTCCGCGCGGGG-----	32979
Qy	97	ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr	116
Db	32980	-----TCCCTCGACTGGCTCATCACCCCTTCTCCCTCTCCACCTTCCCAACACC	33030
Qy	117	LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAla-----	133
Db	33031	CTCGTCATGGCATCCCGTGTCTCAAGGGGATGATCGCGCGCGCGGGCGCGAGCCGGC	33090
Qy	134	-----GlyThrLeuMetValGlnValValLeuGlnCysIleIleTrpTyrThr	150
Db	33091	GCGAGCTCCGGGACGCTCATGGTGCAGATCTGTCTCCATGTGATCATCTGGTACAG	33150
Qy	151	LeuMetLeuPheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPhePro	170
Db	33151	CTCATGCTTCTCTTCAGTAGTACCCTGGCGCGGCTTCTTGTTCATGGAGCAGTTCGCG	33210
Qy	171	AspGlyAlaAlaAlaSerIleValSerPheArgValAspSerAspValValSerIleAla	190
Db	33211	GAC-----ACCGCCGCGCTCCATCGTCTCTTCGCGGTGGATTCGAGCTGTCTCGCTCGCC	33267
Qy	191	ArgGlyAsp-----ValGluLeuGluAlaGluProAspGlyValAlaGlyIle	207

Db	33268	GGCGGCGTGGTGGAGCCCGCGAGCTCGACGGGAG-----	33303
Qy	208	AlaValSerSerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThr	227
Db	33304	-----GCAGAGGTGGGACCATGCGAAGATCGAGTACCGTGCAGAGTTCGACG	33354
Qy	228	SerSerArgSerGluAlaAlaCysSerHis-----SerHisSerGluThrMetGln	244
Db	33355	AGTCTCGCTCCGAGGGCGTGCTCGACGGGACGCGAGTCACTCGAGTCCATGACG	33414
Qy	245	ProArgValSerAsnLeuSerGlyValGluLeuLysSerLeuGlnSerSerArgAsnPro	264
Db	33415	CGCGGCGTCTCAACCTCTCCGGCTCGAGATTATCTCGCTGCAGTCTCGCGGACCCG	33474
Qy	265	ThrProArgGlySerSerPheAsnHisAlaAspPheAsnIleValGlyAlaAla	284
Db	33475	ACGCGCGGGGCTCCAGCTTCAACCAACCGCGAGTCTTCAACATCGTC-----	33522
Qy	285	LysGlyGlyGlyAlaAlaGlyAspGluGluLysGlyAlaCysGlyGlyGlyGlyGly	304
Db	33523	-----GGCAACGGCAACGACGCGCAGGAGAGAGGGCGCC-----GCGGTGGCGGG	33570
Qy	305	GlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLysAspLysHisMet	324
Db	33571	GGCCACTCCCGCAGCCG-----GFGTGGGAGAGAGAAAGCACTGCACATG	33618
Qy	325	LeuValTyrSerSerSerAlaSerProValSerGluArgAla-----	338
Db	33619	TTCGTGTGAGCTCAAGCGCTCGCGGTGTGGAGCGCGCGCGCGCGCGCTGCC	33678
Qy	339	AlaValHisValPheGlyAlaGlyAlaAspHisAlaAsp-----	352
Db	33679	GGCGCGGTGCAGTCTCTCGCGCGTGTGGCGCGACCGCGCGCGCAAAAGGTGACGAC	33738
Qy	352	-----	352
Db	33739	ACCTTTTATCATCTTTTCGTGCACCTTTGACTTCCTCTTTTGTGTTTGCCCGCGTGATT	33798
Qy	353	-----ValLeu-Ala-----	355
Db	33799	ACACTGTGTTGCTCTGCACAAATAATACACTCTTGTCTGTGCTCTTGACCACTCGAT	33858
Qy	355	-----	355
Db	33859	CAGTATAATCAACGGTCCAGATCACACTGCTGCTGATGTTGTTGGAGAACTCAACTGG	33918
Qy	355	-----	355
Db	33919	TATACTGATTAAATCAACCATCTTGAAGTGATTCTTTTTCACAAACTTACCTTCTTGAA	33978
Qy	355	-----	355
Db	33979	TTGTTAATTATGTAAGCTTTTCTGATCTTAACAGTGCATTGACTGTTAAGTAGAGTTG	34038
Qy	355	-----	355
Db	34039	AATTGTACAGTTAAGCAAAATCTTAAACAGTAAACTACTGATTTCACAAAAGAAAATGTT	34098
Qy	356	-----LysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSer	371
Db	34099	CTTTTGATGCATGCACAGGTGCTCAGCGCTATGATAG-----TACAGCTT	34143
Qy	371	ArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGln	391
Db	34144	CGGGAACAAGAT-----GAGAAGACGGGGCCACACTGTGAGAGCTGGG	34188
Qy	391	YSerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArgAlaAlaAlaVal	411
Db	34189	GTCCAACTCGACGGCGCAGCTCCGGCCAAAGACACGCGCAGGAGGAGGGCGCGCA	34243
Qy	411	IleMetProProAlaSerValMetThrArgLeuIleLeuIleMetValTyrArgLysLeu	431
Db	34244	GGGATCCCGCGCGGAGCGTATACAGAGGTCTACTACTGATCATGTGCTGTGAGAGCT	34302

```

Qy 431 uileArgAnProAanThrTySerSerLeuIleGlyValValTrpSerLeuValSerTy 451
Db 34303 GATCAGGAACCAAACTTACTCCAGCCTCTCGGTGTCATCTGGTCCCTCGTCTCCTA 34362
Qy 451 r----- 451
Db 34363 CAGGTCAATTAAGCAACATTTTACTCACATTTTTCACATGGCCATGTCATCATGTC 34422
Qy 452 -----ArgTrpG1 454
Db 34423 ATCTCTGTTGTGTCATGTTGTTAACTCTGTGAACTATGATGTAATTCAGGTGGG 34482
Qy 454 VileGluMetProAlaIleAlaAArgSerIleSerIleLeuSerAseAlaGlyLeuG1 474
Db 34483 AATTGAGATGCGGTATCATCGCCCGGTGATCTTCGATCTTCAGATGACAGGCTCG 34542
Qy 474 yMetAlaMetPheSerLeu----- 480
Db 34543 AATGGCCATGTTTCAGCCTAGGTTTGTGTCATCTCTGATACATGCTGTCATTCGTAA 34602
Qy 481 -----GlyLeuP 483
Db 34603 TCTTACATGGTTCATGTCGTGTCGAATGACTGACTGATGTTCTGGTTTCAGAGATGT 34662
Qy 483 heMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMetG 503
Db 34663 TCATGGCATTCAGCCAGGATCATGCTGTGGAACTCCCTTGTCTGATGCTATGG 34722
Qy 503 lyValArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuA 523
Db 34723 CGGTCAAGTTCCTGTTGCTGCGGTGATGCTGCGGCTTCATGCGCGCTCGACTTC 34782
Qy 523 rgGlyValLeuLeuHisIleAlaIleVal----- 532
Db 34783 GGGGGTGTCTCTGCACATTCGATGTTTCAGGTAAGCTGCAAGTTCAGGCTTCGACC 34842
Qy 533 -----G 533
Db 34843 AAAATGCATCCTTAAGTCAAGAAAGAAATCTATGTAATCAACAATGTCATGACTTC 34902
Qy 533 InAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlyTrpGlyValHisP 553
Db 34903 AGCGCGCTTCTCTCAAGGAATCGTGGCTTCGTGTTGGCAAGAGATACAATGTTTCATC 34962
Qy 553 roAspIleLeuSerThrAlaTyTrpGlyProIleThrSer 565
Db 34963 CTAACATTCAGCACAGCGTATGATCTCTTTACCACC 35000

RESULT 6
AC123528/c 141025 bp DNA linear HTG 29-MAY-2002
LOCUS Oryza sativa chromosome 11 clone OSUNBa0078E03, *** SEQUENCING IN
DEFINITION PROGRESS ***, 3 ordered pieces.
ACCESSION AC123528
VERSION AC123528.1 GI:21240714
KEYWORDS HTG: HTGS, PHASE2.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
REFERENCE 1 (bases 1 to 141025)
AUTHORS Linton, E.W., Tanyolac, B., Young, S., Kavchok, S., Keizer, G.,
Bronzino, A., and Messing, J.
TITLE Unpublished
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 141025)
AUTHORS Linton, E.W., Tanyolac, B., Young, S., Kavchok, S., Keizer, G.,
Bronzino, A., and Messing, J.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) The Plant Genome Initiative at Rutgers -
Waksman Institute, Rutgers, The State University of New Jersey, 190

```

Prelinghuysen Road, Piscataway, New Jersey 08854, USA  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 3 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submittor.

\* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 12796: contig of 12796 bp in length  
 \* 12797 12896: gap of unknown length  
 \* 12897 109308: contig of 96412 bp in length  
 \* 109309 109408: gap of unknown length  
 \* 109409 141025: contig of 31617 bp in length.

FEATURES  
 source  
 1..141025  
 /organism="Oryza sativa"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:4530"  
 /chromosome="11"  
 /clone="OSUNBa0078E03"

## ORIGIN

Alignment Scores:  
 Pred. No.: 2,01e-99 Length: 141025  
 Score: 1959.50 Matches: 456  
 Percent Similarity: 58.92% Conservative: 26  
 Best Local Similarity: 55.75% Mismatches: 43  
 Query Match: 67.34% Indels: 294  
 DB: 2 Gaps: 19

US-10-030-884-14 (1-573) x AC123528 (1-141025)

```

Qy 1 MetIleThrAlaLeuAspLeuTyHisValLeuThrAlaValProLeuTyValAla 20
Db 62955 ATGATCAGCGTGTGACCTGTACCACTGCTGACGCGGTGTGCGGTGTCGCGG 62896
Qy 21 MetThrLeuAlaTyGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40
Db 62895 ATGACGTGCGGTACCGCTCCGTCGGTGTGGCGCATTTTCTCCCGACGAGTCTCC 62836
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheLeuSer 60
Db 62835 GGCATCAACCGCTTCGTCGCGCTTCTGCGCGTCTCTCTCTCTCTCTCTCTCTCT 62776
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
Db 62775 ACCAACAACCCCTTCGCCATGAACCTCGCTTCTCGCGCGCACACGTTCCGAGAGCTC 62716
Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96
Db 62715 ATCGTCTCGCCCTCTCTCGCCCTCTGTCGCGGCTCTCGCGCGCGG----- 62668
Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116
Db 62667 -----TCACTCGACTGGCTCATCACCTCTTCTCTCTCTCTCTCTCTCTCTCTCT 62617
Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyGlyAlaSerSer-----Ala 133
Db 62616 CTGTCTATGGGCTATCCGCTGTCTCAAGGGGATGTACGCCGCCGCCGCCGCGCGTCACTCC 62557
Qy 134 GlyThrLeuMetValGlnValValValValLeuGlnCysIleIleTrpTrpThrLeuMetLeu 153
Db 62556 GGCAGCCTCATGTTGTCAGATCGTCGTCTCCATGTCATCATCTGTGTACACGCTCATGCTG 62497
Qy 154 PheIlePheGluTyTrpArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAla 173
Db 62496 TTCTCTTTCAGTACCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 62440
Qy 174 AlaAlaSerIleValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAsp 193
Db 62439 GCGGCTTCCATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 62380

```



REFERENCE 2 (bases 1 to 2297)  
 AUTHORS Baba,K., Schrader,J., Palme,K., Bhalerao,R.P. and Sandberg,G.  
 TITLE A family of PIN1 like auxin transporters in hybrid aspen  
 JOURNAL Unpublished

REFERENCE 3 (bases 1 to 2297)  
 AUTHORS Baba,K., Schrader,J., Palme,K., Bhalerao,R.P. and Sandberg,G.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-MAY-2002) Forest Genetics and Plant Physiology,  
 Swedish University of Agricultural Sciences, Umea 90183, Sweden

FEATURES  
 source  
 1..2297  
 /organism="Populus tremula x Populus tremuloides"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:47664"  
 1..2297  
 /gene="pin2"  
 339..2105  
 /gene="pin2"  
 /codon\_start=1  
 /product="PIN1-like auxin transport protein"  
 /protein\_id="AA054034.1"  
 /db\_xref="GI:21435940"  
 /translation="MISIVDLVHLTVAVPLVAVMLLAGVSKWKIFSPDQCSGINSR  
 FVALFAPLISLHFPISNNPYAMNGLFIADILQKILVILVLAISWRASSRGSLEMSI  
 TLFSSLSPTLVMLGIPLLKLGNSGSLMVOIVLQCIWITMLFETRGARIL  
 IGEQFDTAGSIISPRVDIILSDREPLEQAEVGEDKLHVTVRKSTSRSRVFS  
 HMSGNLGSLTPRSDNLNIAIYLSQSRNPTPRASFNHTDYFVWNGKNSPRH  
 SNFTLPQBPESGGLVFGNVPNGSAVPAFPNAGIFSPGGKKKANGKGLHMF  
 VNSSASPVSEGLHVFEGKDGNDLGGVAHKHDVEFGDRDFSGNRPNGNVDRDG  
 PVLKLGSSSTAEHLPKSANGAEKPTWMPPTVMTRLLILIMVSKLIRNPNVYSLLI  
 GLWLSVSPFKNVMPATIANISAILNSAGLGMKMFSLGLFWALPRIIACGNSIASF  
 AMGVRLETPGAVMAASFAVGLRGDLRLIAIVQAAIPQGVFPVFAKGINVHPDILST  
 GVIFGMLIALPITLVYVILLGL"

ORIGIN

Alignment Scores:  
 Pred. No.: 8 51e-92 Length: 2297  
 Score: 1785.50 Matches: 384  
 Percent Similarity: 71.94% Conservative: 57  
 Best Local Similarity: 62.64% Mismatches: 89  
 Query Match: 61.36% Indels: 83  
 DB: 8 Gaps: 16

US-10-030-884-14 (1-573) x AF515435 (1-2297)

Qy 1 MetIleThrAlaLeuAspLeuTyHisValLeuThrAlaValProLeuTyHisValAla 20  
 Db 339 ATGATCATGATCGTAGACCTCTACCATGCTTACAGCTGTGGTACCCCTCTATGTGGCC 398  
 Qy 21 MetThrLeuAlaTyGlySerValArgTTPArgIlePheThrProAspGlnCysSer 40  
 Db 399 ATGATTTGGCTTAGTTCCTGTCATGATGTTGGAATCTTTAGCCCTGATCAATGTTC 458  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 459 GGGATCAACAGATTGTGCTGTGTGCAGTACCTTGTCTTTCTTTTCACTTCATTCC 518  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 519 TCCAAACACCCCTATGCTAGAACTAGGTTCACTGCTGCATGATCTTTCAAAAAT 578  
 Qy 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 Db 579 ATTGTCTTGGTGTATTAGTATTGTGCCAGA-----GCTAGCTCCAGAGGC---TCC 629  
 Qy 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrIleValMetGly 120  
 Db 630 CTTGAATGGTCCATTACTTTGTTTCACTCTCTAGTCTCCCAACACTCTTGTATGGGC 689  
 Qy 121 IleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 Db 690 ATCCCTTGTGTAAGGGAATGATGGA---AATTCCTCAGGAGTCTCATGTGCCAATA 746

Qy 141 ValValLeuGlnCysIleIleTTPTrThrLeuMetLeuPheLeuPheGluTyArgAla 160  
 Db 747 GTTGTCTTCAATGCATATATATGTCATCATGCTGCTTCTTGTGTAGTATAGGA 806  
 Qy 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 Db 807 GCTAGATCTTGAATGGTGAACAGATTTCCTGAC---ACAGCTGGCTCTATAATATCTTC 863  
 Qy 181 ArgValAspSerAspValValSerLeu---AlaArgGlyAspValGluLeuGluAlaGlu 199  
 Db 864 AGGTTGATCTCATATTCTTTCTTTAGATGGTAGAGGCCATTGCAACCCAGGCTGAA 923  
 Qy 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyClyAspAlaGlyArgVal 219  
 Db 924 GTC-----GGTGAAGATGGGAAGTTA 944  
 Qy 220 ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis 239  
 Db 945 CATGTTACTGTTCCGAATCACTAGTTCAGATCAGAGTCTTTCTCATCATGCTCCAT 1004  
 Qy 240 -----SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIle 255  
 Db 1005 GGTCTAAATCTGTCTTTCATTGACTCTCTAGACCATCTAATTTGACCAATGCAGAGATA 1064  
 Qy 256 TyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAsp 275  
 Db 1065 TACTCTCTACAACTCCTCAAGAAATCCTACCCCAAGAGCTTCGAGTTTAAACCATACCGAT 1124  
 Qy 276 PhePheAsnIleValGlyAlaAlaLysGlyGlyGlyAlaAlaGly----- 292  
 Db 1125 TTTTACTCCATGGTT-----AATGGCAAGAATCGAGTCCACAGACACTCT 1169  
 Qy 293 -----AspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyGly 305  
 Db 1170 AACTTCACTAATTGCAATTCGACGAAGAAGTGA-----GACTTGGGTATTGGT 1223  
 Qy 306 HisSerProGlnProGlnAlaValAlaValProAla----- 317  
 Db 1224 AATGTTCCAGAGCAATGGAAGTCTTATCTCTCTCCACCAATGCTGGAATCTTTCT 1283  
 Qy 318 -----LysArgLysAspLeuHisMetLeuVal 326  
 Db 1284 CTGGGGGTAAAGAGGCAAAATGGGGCTGAGAAATGGCAAGGATTTGCATATGTTGTT 1343  
 Qy 327 TrpSerSerAlaSerProValSerGluArgAlaAlaValHisValPheGlyAlaGly 346  
 Db 1344 TGGAGTTCAAGTGTTCACCAAGTATCTGAA---GGTGGCTACATGCTTTAAGGGAGG 1400  
 Qy 347 Gly-----AlaAspHisAlaAspValLeuAlaLysGlyAla 358  
 Db 1401 GATTATGGCAATGACCTCGGTGGGTAGCTCACCATAAAGAT----- 1442  
 Qy 359 GlnAlaTyArgGluTyArgAspAspTySerSerArgThrLysAsnGlySerGly 378  
 Db 1443 -----TATGATGATGTTGGTTCGAGATGAGTTCAGCTTGGAAATAGACCTGGACCTAAT 1496  
 Qy 379 GlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeu 398  
 Db 1497 GGAGTGGATCGGGATGTCTCAGTCTTTTAAAGTGTGTTCAAGCTCTACAGCCGAGCTA 1556  
 Qy 399 TyrProLysAspAspGlyGluGlyArgAlaAlaValAlaMetProAlaSerVal 418  
 Db 1557 CACCCAAAGAGTGTCTGTAATGGTGAAGCCAGCCAACTGTCTCTCTACAGCGTT 1616  
 Qy 419 MetThrArgLeuIleLeuMetValTrpArgLysLeuIleArgAsnProAsnThrTyr 438  
 Db 1617 ATGACAAGACTCACTCTCATTTATGTTGGAGAAAATCTTATAGAAAATCCAAATACCTAT 1676  
 Qy 439 SerSerLeuIleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetPro 458  
 Db 1677 TCCAGCTGATGGCCCTCCTCTGTTCTTCTTTCATTCAGTGGAAACATATGATGCC 1736  
 Qy 459 AlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPhe 478





Itch, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
 Kishikawa-Hirzane, T., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,  
 Koya, S., Kurinara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numaaki, R., Ohno, M., Osato, N.,  
 Ota, Y., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, J., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
 Takaku-Akashira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
 Yasunishi, A. and Hayashizaki, Y.

## FEATURES

Location/Qualifiers  
 1..2402  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /culti\_var="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J033091B16"

## ORIGIN

## Alignment Scores:

Pred. No.: 5,56e-85 Length: 2402  
 Score: 1665.00 Matches: 365  
 Percent Similarity: 67.58% Conservative: 54  
 Best Local Similarity: 58.87% Mismatches: 111  
 Query Match: 57.22% Indels: 90  
 DB: 8 Gaps: 17

US-10-030-884-14 (1-573) x AK102343 (1-2402)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
 Db 151 ATGATTACGGCGCGGAGCTTCTTACCACGTGATGACGGCGATGGTGGCGGTGACGTGGCG 210  
 Qy 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40  
 Db 211 ATGATACGTGGCGTACGGGTGGTGAAGTGGTGGCGCATCTTTCACGCCGACGATGTCTCC 270  
 Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 271 GGGATCAACCGCTTCGTGGCGCTCTTCGCGGTGGCGTGGTGGTGGTGGTGGTGGTGGTGG 330  
 Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAspThrLeuGlnIysVal 80  
 Db 331 ACCAACACCCGGTACACGATGAACCTCCGGTTCATCGCGCGGACACGCTCGAAGCTG 390  
 Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
 Db 391 ATGGTGTGGCCATCTCACGGCTGGAGCCACCTCAGCCCGCGGGG-----438  
 Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 439 -----AGCCTCGAGTGGACCATCATCAGCTCTTCTCCCTCTCCACGCTGCCAACACG 489  
 Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
 Db 490 CTCGTCATGGGATCCCTTGGCTCAAGGCATGATACGGGGAGTCTCC---GGCAGCTC 546  
 Qy 137 MetValGlnValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPhe 156  
 Db 547 ATGGTGCAGATCGTGTGTGCAGTGCATCATCTGGTACACGCTCATGCTCTTCATGTC 606  
 Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
 Db 607 GAGTACCGCGGCGCGCGATGTCATACCGAGCATTCCTCCGAGC---ACCGCGCCCAAC 663  
 Qy 177 IleValSerPheArgValAspSerAspValValSerLeu-----AlaArgGlyAspVal 194  
 Db 664 ATCGCTCATCGTGTGACCGCGACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 723  
 Qy 195 GluLeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGly 214  
 Db 724 GAGACGGAGACGAGGTGAAG-----744  
 Qy 215 AspAlaGlyArgValArgValThrValArgIysSerThrSerSerArgSerGluAlaAla 234

Db 745 GAGGACGGGAGGATACACGTCACCGTGGCCGCTCCAAACGCGTCTGCTCGGACATCTAC 804  
 Qy 235 CysSerHisSerHisSer---GlnThrMetGlnProArgValSerAsnLeuSerGlyVal 253  
 Db 805 TCCCGCGCTCCATGGCTTCTCCAGCACACCGCGCGCGGCAACCTCACCAACGCC 864  
 Qy 254 GluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHis 273  
 Db 865 GAGATCTACTCGTCTCGTCGCGGAAACCGGACCGCGGAGGGTTCAAGCTTCAACACC 924  
 Qy 274 AlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAsp 293  
 Db 925 ACCGACTTCTACTCCATGTGGCGCGCAGCTCCAACCTCGGCGCGCCACGCGTTCGGC 984  
 Qy 294 GluGluLysGlyAlaCys-----299  
 Db 985 GTCCGACCGCGGCCACGCGCGCGCTCCAACTACGAGGACGACGCTCAAGCCCAAG 1044  
 Qy 300 -----GlyGlyGlyGlyGlyGlyHis-----SerProGlnPro 310  
 Db 1045 TACCCGCTCCGGCGTGCATGGCGGCCCATGGCGGCATATCCCGGCGCGACCCG 1104  
 Qy 311 GlnAlaValAlaValPro-----AlaLysArg 319  
 Db 1105 GCGGTCTCTGTCGGCGCCCAAGAGCGCGCCAAAGCGCGCCACGAACGGCGAGCCAAAGGC 1164  
 Qy 320 LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAla 339  
 Db 1165 GAGGACCTCCATGTTCTGTGGAGCTCCAGCGCTCGCGCTGCCGCTGTCGAC-----1215  
 Qy 340 ValHisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAla-----355  
 Db 1216 -----GTCTTCGGCGCGCGCGCCAGACTACAACGCGCGCGCGCTCAAGTCCCCC 1269  
 Qy 356 -----LysGlyAlaGlnAlaTyrAspGluTyr---GlyArgAspAspTyrSerSer 371  
 Db 1270 CGCAAAATCGATGGACGAGGACGAGGAGGACTACGTGGAGCGGGACCATTTTCAGCTTC 1329  
 Qy 372 ArgThrLysAsnGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGly 391  
 Db 1330 GGGAAACAGG-----GGCGTCATGGACAGGAC-----CGGAGGCGAGG 1368  
 Qy 392 SerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGlyArgAlaAlaAlaVal 411  
 Db 1369 GACGAGAGCG 1428  
 Qy 412 AlaMetProProAlaSerValMetThrArgLeuLeuLeuMetValTrpArgLysLeu 431  
 Db 1429 GCGATCCGCGCGACGAGCGTGTATGACCCGCTCATCTGATCATGTGTGTGGCGACGCTC 1488  
 Qy 432 IleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSerLeuValSerTyr 451  
 Db 1489 ATCCGCAACCGCAACCTACTCCAGCTCATCGGCTCATCTGGTCCCTCGTCTGCTTC 1548  
 Qy 452 ArgTyrGlyIleGluMetProAlaIleIleAlaArgSerIleSerIleLeuSerAspAla 471  
 Db 1549 AGGTGAACATTCAGATGCCGCGCATGTCCTGCTGAATCATCTCTGATCTCTGTCGAGCGG 1608  
 Qy 472 GlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIle 491  
 Db 1609 GGGCTCGGCATGGCATGTTCACTCTCGTCTGTTTCATGGCTGCAGCGCGACATCATC 1668  
 Qy 492 AlaCysGlyAsnLysLeuAlaIleAlaMetGlyValArgPheValAlaGlyProAla 511  
 Db 1669 GCGTGGGGAACAAAGTGGCGAGTACGCGCATGGCGGTGCGGTCTCTGGCGCGCGCGG 1728  
 Qy 512 ValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIle 531  
 Db 1729 GTGATGGCGCGCGCTCTCTCGCGCTCGGACTCCGTGGCGACGCTCTCTGACGCTGCA 1788  
 Qy 532 ValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGluTyrGlyVal 551

Db 1789 GTCCAGGCGAGCTCTCCCGCCAGGCGCATTCCTTCCTTCGCGGAGGAGTACAGCGTG 1848

QY 552 HisProAspLeuSerThrAla-----TyrGly linear PLN 24-JUL-2003  
|||||  
1849 CACCCTAGCATCTCAGCAGCAGCTGTCATCTTGGCATGCTCATCGCTGCGCTATCAC 1908

RESULT 10  
AK101504  
LOCUS  
DEFINITION  
Oriza sativa (japonica cultivar-group) cDNA clone:J033044E23, full  
insert sequence.  
AK101504  
GI:32986713  
VERSION  
FLI CDNA: CAP trapper  
KEYWORDS  
SOURCE  
ORGANISM  
Oriza sativa (japonica cultivar-group)  
Oriza sativa (japonica cultivar-group)  
Spermatophyta; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE  
AUTHORS  
1 The Rice Full-length cDNA Consortium, National Institute of  
Agrobiological Sciences Rice Full-length cDNA Project Team,  
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,  
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
Science Genome Sequencing & Analysis Group: Ohtsuki, Y., Murakami, K.,  
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,  
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,  
Kawai, J., Carninci, P., Hayatsu, N., Imotani, K., Shiraki, T.,  
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Ito, M.,  
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oota, Y.,  
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M. and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from  
japonica rice  
Science 301 (5631), 376-379 (2003)  
22752273  
12869764  
2 (bases 1 to 2457)

TITLE  
JOURNAL  
MEDLINE  
PUBMED  
REFERENCE  
AUTHORS  
1 MetileThraAlaLeuAspLeuTyHHisValLeuThraAlaValProLeuTyValala 20  
352 ATGATATCCGGGACACACTTCTACACGGTGTGGGGGGGGTGGCGGTGTACGTGGCG 411  
QY 21 MetThreLeuAlaTyGlySerValArgTrrTrrPrrArgPheThrProAspGlnCysSer 40  
412 ATGTTCCTGGCGTACCGGTTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 471  
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
472 GGATATCAACCGCTTCTGCGCCATCTTTCGCGGTGGCGGTGGCGGTGGCGGTGGCG 531  
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThreLeuGlnVal 80  
532 ACCAAGACCGGTACGCCATGAACCTTCGCTTCTGGGGGGGGGACACACGTGCGAAGCTG 591  
QY 81 AlaValLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuGly 100  
592 CTGCTCTGGCGGGGTCTCGCGCGGTGGTGGCGGTGGTGGCGGTGGCGGTGGCGGTGG 651  
QY 101 LeuAspTrpSerIleThreLeuPheSerLeuSerThreLeuProAsnThreLeuValMetGly 120  
652 CTGGAGCTGGTTCATCAGCTCTTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 711  
QY 121 IleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThreLeuMetValGlnVal 140  
712 ATCCCGCTGCTGATGCCCATGTACGGGCCACTACTCC---GGCTCGCTCATGTTCAGATC 768  
QY 141 ValValLeuGlnCysIleIleIleTyrrThreLeuMetLeuPheLeuGluTyrrArgala 160

URL : <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,  
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and  
Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Ohtsuki, Y., Iida, Y.,  
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,  
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center  
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishi, Y.,  
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oota, N.,  
Ota, Y., Saitoh, H., Sakai, K., Sakai, K., Sakazume, N., Sano, H.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
Yasunishi, A. and Hayashizaki, Y.

FEATURES  
Location/Qualifiers  
source  
1..2457  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033044E23"

ORIGIN  
Alignment Scores:  
Pred. No.: 1-17e-84 Length: 2457  
Score: 1659.50 Matches: 378  
Percent Similarity: 65.99% Conservative: 47  
Best local Similarity: 58.70% Mismatches: 104  
Query Match: 57.03% Indels: 115  
DB: 8 Gaps: 18

US-10-030-884-14 (1-573) x AK101504 (1-2457)

QY 1 MetileThraAlaLeuAspLeuTyHHisValLeuThraAlaValProLeuTyValala 20  
Db 352 ATGATATCCGGGACACACTTCTACACGGTGTGGGGGGGGTGGCGGTGTACGTGGCG 411  
QY 21 MetThreLeuAlaTyGlySerValArgTrrTrrPrrArgPheThrProAspGlnCysSer 40  
Db 412 ATGTTCCTGGCGTACCGGTTCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 471  
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 472 GGATATCAACCGCTTCTGCGCCATCTTTCGCGGTGGCGGTGGCGGTGGCGGTGGCG 531  
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThreLeuGlnVal 80  
Db 532 ACCAAGACCGGTACGCCATGAACCTTCGCTTCTGGGGGGGGGACACACGTGCGAAGCTG 591  
QY 81 AlaValLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuGly 100  
Db 592 CTGCTCTGGCGGGGTCTCGCGCGGTGGTGGCGGTGGTGGCGGTGGCGGTGGCGGTGG 651  
QY 101 LeuAspTrpSerIleThreLeuPheSerLeuSerThreLeuProAsnThreLeuValMetGly 120  
Db 652 CTGGAGCTGGTTCATCAGCTCTTCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCT 711  
QY 121 IleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThreLeuMetValGlnVal 140  
Db 712 ATCCCGCTGCTGATGCCCATGTACGGGCCACTACTCC---GGCTCGCTCATGTTCAGATC 768  
QY 141 ValValLeuGlnCysIleIleIleTyrrThreLeuMetLeuPheLeuGluTyrrArgala 160

Db 769 GTCTGTCCTCCAGTGCATCATCTGGTACACGCTGATGCTCTTCCACTTCGAGTTCCTCCGCC 828  
 Qy 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 Db 829 GCGGGAAGCTGATCGCGACGACAGTTTCGGAC--ACGGGGGGTCCATCGTGTCTCCCTG 885  
 Qy 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
 Db 886 CACGTTCGACCGCGAGTGTGTCTGCTGAGGGCGGCCACGCGGAGACGGAGCGCGAG-- 942  
 Qy 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
 Db 943 -----GTGGGGGG-----GACGGCGGCTGCAC 966  
 Qy 221 ValThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSerHisSer 240  
 Db 967 GTCACTGCGCGCGCTCTCGGTGTCGGG-----CGGTGCG 1002  
 Qy 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260  
 Db 1003 CTGCTGGTTCACGCCGCGCGCTCGAACCTGACGGAGCGGAGATCTACTCGCTTAGTCTCG 1062  
 Qy 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIleVal 280  
 Db 1063 TCGCGAACCACCAACCCCGCGGCTCCCACTTCAACCAACCGCGACTTCTTCGCCATGTC 1122  
 Qy 281 Gly-----AlaAlaAlaLysGlyGlyGlyGlyGlyAlaAlaGly 292  
 Db 1123 GCGCGCGCGGCCACCCCGCCCGACCGCTCGGTGCGCGCTCGAGCTTCGCGCGCTCC 1182  
 Qy 293 Asp-----GluGlu 295  
 Db 1183 GAGCTCTACTCGCTGCAATCTGCGGGGGCCCAACCCGAGGCGAGTCCAACTTCGACGAG 1242  
 Qy 296 LysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaVal 315  
 Db 1243 CACTCGGCA-----CGGCCACCGAAGAACCCACCGCGCAACGACCCACG 1281  
 Qy 316 ProAla-----LysArgLysAspLeuHisMetLeuValTrpSerSerAlaSer 332  
 Db 1282 GGGGCACTCAACGACGATCGCAGAGCTCCCAATGTCGTGTGAGTTCGAGCGGTAT 1341  
 Qy 333 ProValSerGluArgAlaAlaValHisValPhe-----GlyAlaGlyGlyAlaAspHisAla 351  
 Db 1342 CCGGTCTCAGAAGTCAGCGCGCTGCTGTGTTCACTGCGCGCGCGCGCGCGCTCTC 1401  
 Qy 352 AspValLeuAlaLys----- 356  
 Db 1402 GACGTGCGCGCCCAAGAAATCCACATGTCATCCCGCGACCTCCCGCAGAACACCGGC 1461  
 Qy 357 GlyAlaGlnAlaTyrAspGluTyrGlyArgAspTyrSerSerArgThrLysAsnGly 376  
 Db 1462 TCAGGCAAGACGACGAGGAGTACGGC-----GCAGTGGCATTTGGGTGGC 1506  
 Qy 377 SerGlyGlyAlaAspLys-----GlyGlyProThr----- 386  
 Db 1507 GCGCGCGCGGAGAGAACTTCAGTTCGGAGGCGGCAAGAGCGGTGGAGCGCGCGAGGCA 1566  
 Qy 387 -----LeuSerLysLeuGlySerAsnSerThrAla 396  
 Db 1567 GTAGACGAGAGCGCGCTTGTCTGACGGCTGACGAGATGGGTTCGAGTTCGACGCGG 1626  
 Qy 397 GlnLeuTyrProLys-----AspAspGlyGluGlyArgAlaAla 409  
 Db 1627 GAGTGCACCCGAAGGTCTGACGCTGCGACGCGGACCGAGCGCGCGCGCGCGCGCGG 1686  
 Qy 410 Ala-----ValAlaMetProAlaSerValMetThrArgLeuIleLeuMetVal 427  
 Db 1687 GCGGGCGCATCCAAATGCGCGCGGCGGAGTGAACAGCGCTCATCTCATATGATGGT 1746  
 Qy 428 TrpArgLysLeuIleArgAsnProAnthrThrTyrSerSerLeuIleGlyValValTrpSer 447

Db 1747 TGGCGGCAAGCTCATCTCGGAACCCCAACACTTACTCAGCGCTCTCGGCCCTCGCTGGTCC 1806  
 Qy 448 LeuValSerTyrArgTrpGlyIleGluMetProAlaIleIleAlaArgSerIleSerIle 467  
 Db 1807 CTGCTCGCTCTCGGTGGCACGCTCTCCATCGCCAGCAATCGTCGAGAAGTCCATCTCCATT 1866  
 Qy 468 LeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGln 487  
 Db 1867 CTCTCGGACGCGAGCGCTCGGGATGGCCATGTTTAGCTGGGATTTGTTCACTGGCGTGCAG 1926  
 Qy 488 ProArgIleIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheVal 507  
 Db 1927 CCCAGCATCATCGGTGTGGCAATCAGCGCGCTCTCCATCGCGCTCGCGCTTCCTC 1986  
 Qy 508 AlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeu 527  
 Db 1987 GCGGCGCTCGCGCTCATGCGCGCGCTCAATCGCATCGGACTCCGCGGAGCGCTCTCTG 2046  
 Qy 528 HisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLys 547  
 Db 2047 CAGTCCGCTATGTTTCCAGCGCGCTCTACCAAGGATTTGCTGCTTTGTTTTCGAAA 2106  
 Qy 548 GluTyrGlyValHisProAspIleLeuSerThrAla-----TyrGly----- 561  
 Db 2107 GAATACAATGTCCACCGCGCCATCTCGACACAGCGGTAAATTTTGGCATGCTAATAGCT 2166  
 Qy 562 ---ProIleThr 564  
 Db 2167 CTTCGAATCA 2178  
 RESULT 11  
 AK099634  
 LOCUS  
 DEFINITION  
 Oryza sativa (japonica cultivar-group) cDNA clone:J013059P15, full insert sequence.  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa (japonica cultivar-group)  
 Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE  
 AUTHORS  
 The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team:  
 Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ohtomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusunegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Tshii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
 Science 301 (5631), 376-379 (2003)  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 1 (bases 1 to 2343)  
 2 (bases 1 to 2343)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,

Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Negata, I., Nakamura, M., Namiki, T., Narikawa, R., Nishikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otsu, Y., Ryo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyata, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skkuchin@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)  
This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nishikura, J., Oka, M., Ryo, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kangawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Satoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyata, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1. .2343  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J013059F15"

ORIGIN

Alignment Scores:  
Pred. No.: 1,73e-84 Length: 2343  
Score: 1636.00 Matches: 364  
Percent Similarity: 67.42% Conservative: 54  
Best Local Similarity: 58.71% Mismatches: 112  
Query Match: 56.91% Indels: 90  
DB: 8 Gaps: 17

US-10-030-884-14 (1-573) x AK099634 (1-2343)

Qy 1 MetLeuAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
Db 151 ATGATTACGGCGGGGACTTCTACACGTGACGGCGATGGCGGTTGACGGGG 210  
Qy 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40

Db 211 ATGATACTGGCTACGGGTCGGTGAAGTGGTGGCGCATCTTACGGCGGACCAAGTCTCC 270  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 271 GGGATCAACCGTTCGTGGGCTCTTCGGCGTGGCGCTGCTGCTTCCACTTCATCTCC 330  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIysVal 80  
Db 331 ACCAACACCGGTACACGATGAACCTCCGGTTCATCGCGCGGACACGCTGCAGAGCTC 390  
Qy 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
Db 391 ATGGTGTGCCCATGTCTACGGGTCGTGAGGCCACTTCAGCGCGCGGG----- 438  
Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerIleuSerThrLeuProAsnThr 116  
Db 439 -----AGCCTCAGGTGACCATCAGCTCTTCTCCCTCTCCACGCTGCCACACG 489  
Qy 117 LeuValMetGlyIleProLeuLeuAspGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
Db 490 CTGTCTATGGGATCCCTTGTCAAGGCGATGTACGGGAGTTCTCC---GGCAGCCTC 546  
Qy 137 MetValGlnValValValLeuGlnCysIleIleTyrTrpThrLeuMetLeuPheLeuPhe 156  
Db 547 ATGGTCGAGATCGTGTGCTGAGTGCATCATCTGTTACACGCTCATGCTCTTTCATGTT 606  
Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
Db 607 GAGTACCGCGCGCGGATGCTCATCCGCGAGTTTCCCGAC---ACCGCGCGCAAC 663  
Qy 177 IleValSerPheArgValAspSerAspValValSerIleu-----AlaArgGlyAspVal 194  
Db 664 ATGCCTCCATCGTGTGTCGACCGGACGCTGCTGCTGACGCGGAGGAGGACGCATC 723  
Qy 195 GluLeuGluAlaGluProAspGlyValAlaGlyAlaValSerSerArgGlyGly 214  
Db 724 GAGACGGAGCGGAGGTGAAG----- 744  
Qy 215 AspAlaGlyArgValArgValThrValArgIleSerThrSerArgSerGluAlaAla 234  
Db 745 GAGGACGGCAGGATACACGTCACGTCGCGCGGCTCCAAACGCTGCTCGTCGACATCTAC 804  
Qy 235 CysSerHisSerHisSer-----GlnThrMetGlnProArgValSerSerGlyVal 253  
Db 805 TCCCGCGCTCCATGGGCTTCTCCAGCACCGCGCGCGCGGCGGACCACTTCCACCAAGCC 864  
Qy 254 GluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHis 273  
Db 865 GAGATCTACTCGTGTGAGTGTGTCGCGAACCCGACGCGGAGGGTTCAAGCTTCAACCCAC 924  
Qy 274 AlaAspPheAsnIleValGlyAlaAlaAlaIlyGlyGlyGlyAlaAlaGlyAsp 293  
Db 925 ACCGACTTCTACTCCATGGCGCGGACGCTCCAACTTCGCGCGCGCGGCTTCGGC 984  
Qy 294 GluGlyIlyGlyAlaCys----- 299  
Db 985 GTCGCGACCGCGCCACGCGCGCGCTCCAACTACGAGGACGCGCTCCAAAGCCCAAG 1044  
Qy 300 -----GlyGlyGlyGlyGlyHis-----SerProGlnPro 310  
Db 1045 TACCGCTCCCGCGTCCGAATGCGCGCGCCATGTCGCGGCGGCTACCGCGCGCGGAAACCG 1104  
Qy 311 GlnAlaValAlaValPro-----AlaIlyArg 319  
Db 1105 GCGGTGTCTCGCGCGCCCAAGGGCGCCAGNAGCGGCCACGAGCGGCGGCGGCGGCGG 1164  
Qy 320 LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAla 339  
Db 1165 GAGGACCTCCACATGTTCTGCTGAGCTTCCAGCGCGTCCGCGGCTGTCGAC----- 1215  
Qy 340 ValHisValPheGlyAlaGlyIlyAlaAspHisAlaAspValLeuAla----- 355

Db 1216 -----GTCTTTGGCGGGCGGCCAGACTCAACAGCCGCGGAGTCAGTCCCCC 1269

Qy 356 -----LysGlyAlaGlnAlaTyAspLutyr---GlyArgAspAspTyrSerSer 371

Db 1270 CGCAAAATGATGAGCGAGGACGAGGAGGACTACGTGAGCGGAGGAGTTTCAGCTTC 1329

Qy 372 ArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGly 391

Db 1330 GGGAAACAGG-----GGCGTCATGCACAGGAC-----GGGAGGACGAGG 1368

Qy 392 SerAsnSerThrAlaGlnLeuTyProLysAspAspGlyGlyAlaAlaVal 411

Db 1369 GACGAGAAAGCGGGCGGGCGGGCGGCCAGCCAGCAAGCCATGCGGGCGGCGCAGC 1428

Qy 412 AlaMetProProAlaSerValMetThrArgLeuLeuLeuMetValTrpArgLysLeu 431

Db 1429 GCGATGCGCGGCGAGGAGGCGGATGATGACCGGCTCATCTGATCATGGTGTGGCGCAAGCTC 1488

Qy 432 IleArgAsnProAnThrTySerSerLeuLeuLeuGlyValValTrpSerLeuValSerTyr 451

Db 1489 ATCCGCAACCGGAACACCTACTCCAGGCTCATCGGCTCATCTGCTGCTGCTGCTGCT 1548

Qy 452 ArgTTPGlyLeuMetProAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 471

Db 1549 AGGTGGAACTTCGAGATGCGGGCCATGCTCTGAAATCCATCTCGATCTCTCGGACGG 1608

Qy 472 GlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgLysLeu 491

Db 1609 GGGCTCGCATGCGCATGTTTCAGTCTCGTCTGTTTCATGCGTGCAGCGGCACATCATC 1668

Qy 492 AlaCysGlyAsnLysLeuAlaAlaAlaAlaAlaMetGlyValArgPheValAlaGlyProAla 511

Db 1669 GCGTGGCGGAAACAGGTGGCGACGCTAGCATGCGGTGGGTCTCTGGCGGGCGGGCC 1728

Qy 512 ValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIle 531

Db 1729 GTGATGCGCGGGCGGTCTCTCGCGCTCGGACTCGCTGCGGACGCTCTCTGCGACGTCGCCATT 1788

Qy 532 ValGlnAlaAlaLeuProGlnGlyValProPheValPheAlaLysGluTyrGlyVal 551

Db 1789 GTCCAGGAGGCTCTGCCCCAGGAGGATGTTCCCTCTGCTCTGCGCTAGGAGTACAGCGTG 1848

Qy 552 HisProAspIleLeuSerThrAla-----TyrGly-----ProfileThr 564

Db 1849 CACCTAGCATCTTCAGCACAGCTGTGATCTTTGGCATGCTCATCGCTTGCCTATCACC 1908

RESULT 12

AK063976

LOCUS

DEFINITION

AK063976 2126 bp mRNA linear PLN 24-JUL-2003

ORYZA SATIVA (japonica cultivar-group) cDNA clone:001-124-C02, full insert sequence.

ACCESSION

AK063976.1 GI:32973994

VERSION

FLI\_CDNA; oligo-capping.

KEYWORDS

ORYZA SATIVA (japonica cultivar-group)

SOURCE

ORYZA SATIVA (japonica cultivar-group)

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

REFERENCE

AUTHORS

The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team, Kikuchi, S., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shihiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Otono, Y., Sugano, S., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, N., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,

# TITLE JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 2126)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Ikeda, R., Imamura, K., Imoto, K., Itoh, M., Katsuhara, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otono, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shihiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Toyata, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

## Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kanondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: shikichen@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, R., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shihiki, T. and Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otono, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imoto, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugiyama, A., Tanaka, T., Tomaru, A., Toyota, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

## Location/Qualifiers

1. .2126

/organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="001-124-C02"

## FEATURES

source

## ORIGIN

## Alignment Scores:

Pred. No.: 1.64e-84 Length: 2126  
 Score: 1655.50 Matches: 375  
 Percent Similarity: 66.93% Conservative: 42  
 Best Local Similarity: 60.19% Mismatches: 103  
 Query Match: 56.89% Indels: 103  
 DB: 8 Gaps: 17

US-10-030-884-14 (1-573) x AK063976 (1-2126)

QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 DB 108 ATGATATCCGGGACGACTTCTACAGGTGATGGGGCGGTGGTGGTGTACGTGGCG 167  
 QY 21 MetThrLeuAlaTyrGlySerValArgTTPTrpArgIlePheThrProAspGlnCysSer 40  
 DB 168 ATGTTCTTGGCGTACGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 227  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 DB 228 GGATCAACCGCTTCGTCCCATCTTCGGCGTGGCGTCTCTGCTTCCATTCATCTCC 287  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIlyVal 80  
 DB 288 ACCAAGACCGTACGCGCATGACCTCCGCTTCGGCGGCGGACACGCTGCAGAGCTG 347  
 QY 81 AlaValIleAlaLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerProArgAlaLeuGly 100  
 DB 348 CTGCTCTTGGCGGGTTCGGCGGCTGGTTCGGCTCTCCCTCGCGACCGCGCGCGGG 407  
 QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
 DB 408 CTGGACTGGTCCATCATCGCTCTTCTCCCTCTCCACGCTGCCAACACGCTGCTCATGGGG 467  
 QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 DB 468 ATCCCGTCTGTATCGCCATGTACGGGCCATCTCC---GGCTCGCTCATGTTCAGATC 524  
 QY 141 ValValLeuGlnCysIleIleTTPTrpThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
 DB 525 GTCTGTCTCCAGTGCATCTCTGTACAGCTGATGCTTCTCTCTTCAGTTCGCGCGCC 584  
 QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaSerIleValSerPhe 180  
 DB 585 GCAGCGATGTGATCGCGACACGATTCCTCCGAC---ACGCGCGCTCCATCGTGTCTCTG 641  
 QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
 DB 642 CAGCTCGACCGGACGCTGGTGTCTGTGAGGGCGGCACGCGGAGACGCGGCGAG--- 698  
 QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
 DB 699 -----GTGCGCGCG-----GACGGCGCGGTGGAC 722  
 QY 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHisSer 240  
 DB 723 GTACCGGTGGCGGGTCTCGGTGTGGCG-----CGGTGG 758  
 QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260  
 DB 759 CTGTGTGTACCGCGCGCGCGTCCGAACCTGACGGGAGCGGAGATCTACTCGCTTAGCTCG 818  
 QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleVal 280  
 DB 819 TCGCGGAACCAACCCCGCGGGTCTCAACTTCAACCAACCGCGACTTCTTCGCGCATGGTC 878  
 QY 281 Gly-----AlaAlaAlaLysGlyGlyGlyAlaAlaGly 292  
 DB 879 GCGCGGCGGCGCGCGCGCGCGCGCTGCGGTGGCGGCTGAGCTTCGCGCGCTCC 938  
 QY 293 Asp-----GluGlu 295  
 DB 939 GAGCTCTACTCGCTGCAATCGTCTGGGGCCCAACCCGAGGCGAGTCAACTTCGACGAG 998  
 QY 296 LysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 315  
 DB 999 CACTCGGCA-----CGGCCCGCGAAACACCGGCAACGACGACG 1037  
 QY 316 ProAla-----LysArgLysAspLeuHisMetLeuValTrpSerSerSerAlaSer 332  
 DB 1038 GGGGCACTCAACCAACGATGCCAAGAGGCTCCACATGTTCGTGTGGAGTTCGAGCGGTCT 1097  
 QY 333 ProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAsp 352  
 DB 1098 CCGCTCTCAAGATCAAGCTCAGCGCTTCGCTGTGTTTCACTGGCGCGCGCGCGCTTCGAC 1157  
 QY 353 ValLeuAlaLys-----Gly 357  
 DB 1158 GTGGCGCCAAAGAAATCCACATGGTTCATCCCGCGACCTGCGCGAGAACAAACGGGTCA 1217  
 QY 358 AlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySer 377  
 DB 1218 GCGAAGACGACGAGGAGTACGCG-----GCAGTGGCATTTGGGTGGC 1259  
 QY 378 GlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGln 397  
 DB 1260 GCGGCG-----GGCGGAGAACTTCAGC-----TCGACGGCGGAG 1295  
 QY 398 LeuTyrProLys-----AspAspGlyGluGlyArgAlaAlaAla 410  
 DB 1296 CTGACCGCAAGGTGCTGACGCTGACGCGACCGCAACCGCGCGCGCGCGCGCGCGCG 1355  
 QY 411 -----ValAlaMetProProAlaSerValMetThrArgLeuIleLeuIleMetValTrp 428  
 DB 1356 GGGCAGTACCAATGCGCGCGCGGAGGTGATGACAGCTCATCTCTCATTAATGGTGTGG 1415  
 QY 429 ArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSerLeu 448  
 DB 1416 CGCAAGCTCATCGCAACCCCAACACACTTACTTCCAGCTCTCGCGCTCGCTCGCTCTC 1475  
 QY 449 ValSerTyrArgTTPGlyIleGluMetProAlaIleAlaArgSerIleSerIleLeu 468  
 DB 1476 GTCCCTTCGGTGGCGACGCTCCATGCCAGCAATGTCGAGAAGTCCATCTCCATCTC 1535  
 QY 469 SerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnPro 488  
 DB 1536 TCGGACGCGAGCTGGGATGGCCATGTTTAGCTGGGATGTTTCATGGCGCTGACGCC 1595  
 QY 489 ArgIleIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheValAla 508  
 DB 1596 AGCATCATCGCTGTGGCAAAATCAGCGCGCTGCTCTCCATGGCGCTCCGCTTCTCGCG 1655  
 QY 509 GlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHis 528  
 DB 1656 GGCTCTGCGCTCATGGCGCGCGCTCAATCGCATCGGACTCGCGCGGAGCGCTCTCTGAC 1715  
 QY 529 IleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGlu 548  
 DB 1716 GTCGCATTTGTTCAGCGCGCTTACCAACAGGGATGTTGCTTTGTTTTGCAAAAGAA 1775  
 QY 549 TyrGlyValHisProAspIleLeuSerThrAla-----TyrGly----- 561  
 DB 1776 TACATGTCCACCGCGCCATCTTGAGCACAGCGGTAAATTTTGGCATGTCTAATAGCTCT 1835  
 QY 562 ProIleThr 564  
 DB 1836 CCAATCA 1844

## RESULT 13

AK103208

LOCUS

DEFINITION

Oryza sativa (japonica cultivar-group) cDNA clone J033122I23, full

insert sequence.

ACCESSION

AK103208

VERSION

AK103208.1 GI:32988417

AK103208 2470 bp mRNA linear PLN 24-JUL-2003  
 Oryza sativa (japonica cultivar-group) cDNA clone J033122I23, full  
 insert sequence.  
 AK103208  
 AK103208.1 GI:32988417

## KEYWORDS

FLI CDNA; CAP trapper.

## SOURCE

Oryza sativa (japonica cultivar-group)

## ORGANISM

Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

## REFERENCE

## AUTHORS

1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Sugiyama, A., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ootomo, Y., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 2470)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurokawa, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohtsuki, K., Oka, M., Ooka, H., Ootomo, Y., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

Direct Submission

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan [E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007]

This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, K., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohtsuki, K., Oka, M., Ooka, H., Ootomo, Y., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.

Location/Qualifiers

1..2470

source

organism="Oryza sativa (japonica cultivar-group)"

/mol\_type="mRNA"

/cultivar="Nipponbare"

/db\_xref="taxon:39947"

/clone="J033122123"

ORIGIN

Alignment Scores:

Pred. No.: 1,97e-84 Length: 2470

Score: 1655.50 Matches: 377

Percent Similarity: 68.28% Conservatve: 47

Best Local Similarity: 60.71% Mismatches: 103

Query Match: 56.89% Indels: 95

DB: 8 Gaps: 20

US-10-030-884-14 (1-573) x AK103208 (1-2470)

Qy 1 MetIleThrAlaLeuAspLeuThrHisValLeuThrAlaValValProLeuThrValAla 20

Db 139 ATGATTAACGGGGCGGACATTCACACGATGATGCGGATGCGGCTGTACGTGGCG 198

Qy 21 MetThrLeuAlaTyrGlySerValArgTrrpArgIlePheThrProAspGlnCysSer 40

Db 199 ATGATCTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 258

Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheLeuSer 60

Db 259 GGGATCAACCGTTCGTCGGCTCTTCGCGGCTCTTCGCGGCTCTTCGCTTCACATCTC 318

Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnIleVal 80

Db 319 ACCAACAACCCCTACACCATGACCTCCGCTTCATCGCCGCGGACACCTTCGAGAAGCTC 378

Qy 81 AlaValLeuAlaLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100

Db 379 ATCGTCTCGGCTCGTCCCTCACCCTGTGGAGCCAC---CTCTCC-----CGCCGCGGCTCC 429

Qy 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120

Db 430 CTGAGTGGACCATCACCTCTCTCCCTTCACGCTGCCCAACACGCTCGTCATGGGG 489

Qy 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140

Db 490 ATCCCGCTGCTGAAGGGGATGTACGGGGAGTTCTCC---GGTAGCCTCATGTGTGCAGATC 546

Qy 141 ValValLeuGlnCysIleIleIleTrrpThrLeuMetLeuPheLeuPheGluTyrArgAla 160

Db 547 GTGCTGCTCGAGTGCATCATCTCTCTCCCTTCACGCTGCCCAACACGCTCGTCATGGGG 606

Qy 161 AlaArgAlaLeuValLeuAspGlnPheProAsp-----GlyAlaAlaAlaSerIleVal 178

Db 607 GCCAGGATCTCATCCCGGACGATTTCCCGGACACCGCGGGCGGCATCGCTCCATCGTC 666

Qy 179 SerPheArgValAspSerAspValValSerLeu-----AlaArgGlyAspValGluLeu 196

Db 667 -----GTGACCCCGACGCTGCTGCTCGACGCGGGGAGGACATGATCGAGACG 717

Qy 197 GluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAla 216

|||||

Db 718 GAGGCCGAGGTGAG-----CAGGAC 738  
 QY 217 GlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAalaCysSer 236  
 Db 739 GGCAGATACAGTCACTCGGCGGCTCCAAAGCGGTGCGCTCGGAGCTTCTACTCGCGG 798  
 QY 237 HisSerHisSer--GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIle 255  
 Db 799 CGTCTCATGGGTTCTCCAGCACCGCCGCGCGGCGAGCACTCACCAAGCCGAGATC 858  
 QY 256 TyrSerLeuGlnSerSerArgHsnProThrProArgGlySerSerPheAsnHisAlaAsp 275  
 Db 859 TACTCGTCAAGTCTGCGGCAACCGACGCGCGGCGGTCCAGCTTCAACACACTGAC 918  
 QY 276 PhePheAsnLleValGlyAlaAlaLysGlyGlyGlyValAlaAlaGlyAsp----- 293  
 Db 919 TTCTACTCATGTGTCGG-----CGAGCTCAACTTCGCCCGCGGAGCCGCTTC 969  
 QY 294 -----GluGluLysGlyAlaCysGly 300  
 Db 970 GGSGTGGCACCGGCCACCGCGCGCGCTCCAACTACGAGGAGGACGCGCGCGCGCC 1029  
 QY 301 GlyGlyGlyGly-----GlyHisSerProGlnProGln--AlaValAlaValPro 316  
 Db 1030 AACAGCGCGGACGAGTACGCGGCGAGTACCGCGCGGCCAACCCGCGCATGCGCGCGG 1089  
 QY 317 AlaLys-----ArgLysAspLeuHis 323  
 Db 1090 CCCAACCCCAAGAGCGGCCAACGCGGAGCCAGCCAGGCGGAGGAGCGGAAAGCTCCAC 1149  
 QY 324 MetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAlaValHisValPhe 343  
 Db 1150 ATGTTCTGTGGAGCTCCAGCGCGTGGCGCGTGTCCGAC-----GTCTTC 1194  
 QY 344 GlyValGlyValAlaAspHisAlaAspValLeuAla----- 355  
 Db 1195 GGC---AACGGCGCGGAGTACACGACGCGCGCGCTCAAGAGGTCCGATGCGCGTC 1251  
 QY 356 LysGlyAlaAlaAlaTyrAspGluTyrGlyArgAspAspTyrSerSer----- 371  
 Db 1252 GCCTCGCGCGCAAGCGGACGCGGTGGAGAGGAGGACGACTTCAGTTC-CGGGAACAGAGG 1310  
 QY 372 ArgThrLysAsnGlySerGlyGly-AlaAspLysGlyGlyProThrLeuSerLysLeuG 391  
 Db 1311 CGTCGCGGAGGAGCGGAGCGCGGCGGACGAGAGCGGTGGCGCGCGGTGTCGGG 1370  
 QY 391 YSerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGlyArgAlaAlaAlaVal 411  
 Db 1371 TGAGCATGGCAAGCCTGGCTTGACGCG-----GCGCGCGAC 1406  
 QY 411 AlaMetProAlaSerValMetThrArgLeuLleLeuLleMetValTrpArgLysLe 431  
 Db 1407 GCGGATGCGCGCGGACGAGGTGATGACGCGGCTCATCTCGATCATGTTGGTGGCGAAGCT 1466  
 QY 431 uileArgAsnProAsnThrTyrSerSerLeuLleGlyValValTrpSerLeuValSerTy 451  
 Db 1467 CATCCGCAACCCCAACACTTCCAGCTCATCGGCTCATCTGGTCTGCTGCT 1526  
 QY 451 rArgTrpGlyLleGluMetProAlaLleLleAlaArgSerLleSerLleSerAspAl 471  
 Db 1527 CCGTGGAACTTCGAGATGCGCGGATCATCTCTGAAATCATCTCATCTCTCGAGCG 1586  
 QY 471 aGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgLle 491  
 Db 1587 CGGCTCGGATGCGCATGTTTCAGCTTCGCGCTGTTTCATGGCGTGCAGCGCGGATCAT 1646  
 QY 491 eAlaCysGlyAsnLysLeuAlaLleAlaMetGlyValArgPheValAlaGlyProAl 511  
 Db 1647 CGCGTGGCGGAACAAGGTGCGAGCTTCGCGCATGCGGCTGCGGTTCCTGACGCGCGCG 1706  
 QY 511 aValMetAlaAlaSerLleAlaValGlyLeuArgGlyValLeuLeuHisLleAla 531  
 Db 1707 CGTCATGGCGCGGCTTCATGCGCGTCCGCTTCGCGCGGACGCTCTCTCATAGTCCCAT 1766

QY 531 eValGlnAlaAlaLeuProGlnGlyLleValProPheValPheAlaLysGluTyrGlyVal 551  
 Db 1767 CGTGCGAGCGAGCTTGCCACAGGCGATGTCCTTCGCTTCGCGAGAGTACAGCGT 1826  
 QY 551 HisProAspLleLeuSerThrAla-----TyrGly-----ProIleth 564  
 Db 1827 GCACCCCGACATTCCTCAGCAGCGCGGTCACTTTGGCATGCTCATCGCGTGCCCATCAC 1886  
 QY 564 r 564  
 Db 1887 G 1887  
 RESULT 14  
 AX654121 1827 bp DNA linear PAT 22-MAR-2003  
 LOCUS AX654121  
 DEFINITION Sequence 3991 from Patent WO03000898.  
 ACCESSION AX654121  
 VERSION AX654121.1 GI:29156935  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Oryza sativa  
 Oryza sativa  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE  
 AUTHORS  
 Chang,H.S., Chen,W., Cooper,B., Glazebrook,J., Goff,S.A., Hou,Y.M.,  
 Katagiri,F., Quan,S., Tao,Y., Whitam,S., Xie,Z., Zhu,T. and Zou,G.  
 TITLE  
 Plant genes involved in defense against pathogens  
 JOURNAL  
 Patent: WO 03000898-A 3991 03-JAN-2003;  
 Syngenta Participations AG (CH)  
 FEATURES  
 Location/Qualifiers  
 source  
 1..1827  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4530"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 3,86e-84 Length: 1827  
 Score: 1647.50 Matches: 374  
 Percent Similarity: 66.09% Conservative: 45  
 Best Local Similarity: 58.99% Mismatches: 100  
 Query Match: 56.62% Indels: 115  
 DB: 6 Gaps: 18  
 US-10-030-884-14 (1-573) x AX654121 (1-1827)  
 QY 11 LeuThrAlaValValProLeuTyrValAlaMetThrLeuAlaTyrGlySerValArgTrp 30  
 Db 1 ATGGCGCGGTGGTGGCGGTGTACGTGGCGATGTTCTGGCGTACGGGTGCGGTGG 60  
 QY 31 TrpArgLlePheThrProAspGlnCysSerGlyLleAsnArgPheValAlaLeuPheAla 50  
 Db 61 TGGGGCATCTTCACGCGGACCAAGTGTCCGGCATCAACCGCTTCGTGCGCATCTTGCC 120  
 QY 51 ValProLeuLeuSerPheHisPheIleSerThrAsnAspPropheAlaMetAsnLeuArg 70  
 Db 121 GTGCGGTCTCTGCTCTTCCATCTCCACCAACGACCGGTACGCGCATGAACCTCCGC 180  
 QY 71 PheLeuAlaAlaAspThrLeuGlnLysValAlaValLeuAlaLeuAlaLeuAlaSer 90  
 Db 181 TTCTGGCGGCGGACACGCTGCAGAGAGTGTCTGCTGCGGGGCTCGCGCGGTGCTG 240  
 QY 91 ArgGlyLeuSerSerProArgAlaLeuGlyLeuAspTrpSerLleThrLeuPheSerLeu 110  
 Db 241 CGCTCTCCCTCGGAGCGCGCGCGGTGGACTGGTCCATCCAGCTCTTCTCCCTC 300  
 QY 111 SerThrLeuProAsnThrLeuValMetGlyLleProLeuLeuArgGlyMetTyrGlyAla 130  
 Db 301 TCCAGCTGCCCAACACGCTGTATGGGATCCCGTGTGTATCGCATGTAGCGGCCA 360  
 QY 131 SerSerAlaGlyThrLeuMetValGlnValValLeuGlnCysLleIleIleTrpTyrThr 150

```

Db 361 TACTCC---GGTCGCTCATGTCAGATCGTCGTCCTCCAGTCATCATCTGCTACACG 417
Qy 151 LeuMetLeuPheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPhePro 170
Db 418 CTGATGCTCTTCTTTGAGTTCCGCGCGCGCGGATGCTGATCGCCAGCAGTTCCCG 477
Qy 171 AspGlyAlaAlaAspSerLeuValSerPheArgValAspSerAspValValSerLeuAla 190
Db 478 GAC---ACGGCGCGCTCCATCGTCTCCCTGTCACGTCGACCCGGACGTCGTGCTGCG 534
Qy 191 ArgGlyAspValGluLeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSer 210
Db 535 GCGCGCCACCGGACGAGCGGCGGAG---GTGCGCGCG--- 570
Qy 211 SerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThrSerSerArg 230
Db 571 -----GACGGCGGCTGCACGTCACCGTGGCGCGCTCTCGGTGTGCGGG 615
Qy 231 SerGluAlaAlaCysSerHisSerHisSerGlnThrMetGlnProArgValSerAsnLeu 250
Db 616 -----CGGTGCTGCTGCTGTCACCGCGCGCGCGCTCGAACCTG 651
Qy 251 SerGlyValGluIleTyrSerLeuGlnSerArgAsnProThrProArgGlySerSer 270
Db 652 ACGGAGCGGAGATCTACTCGTTAGCTGCTCGCGAACCACACCCCGGGGGCTCCAC 711
Qy 271 PheAsnHisAlaAspPheAsnLeuValGly-----Ala 282
Db 712 TCAACACCGCGGACTTCTTCGCCATGTCGCGCGCGCGCGCACCCCGCGCGCGCT 771
Qy 283 AlaAlaLysGlyGlyGlyAlaAlaGlyAsp----- 293
Db 772 GCGGTGCGCGGCTCGAGCTTCGCGCTCCGAGCTCTACTCGCTCGAATCTGCGCGGGC 831
Qy 294 -----GluGluLysGlyAlaCysGlyGlyGlyGlyGly 305
Db 832 CCAACCCCGAGCGAGTCCAACTTCGAGGAGCACTCGCA----- 870
Qy 306 HisSerProGlnProGlnAlaValAlaValProAla-----LysArgLysAspLeu 322
Db 871 CGCGCCCGGAAACCCCGGCAACGACCGCGGGGCACTCAACGAGTCCCAAGAGTCTC 930
Qy 323 HisMetLeuValTyrSerSerSerAlaSerProValSerGluArgAlaAlaValHisVal 342
Db 931 CACATGTTCTGTGTGAGCTCGAGCGCTCTCCGCTCTCAGAGTCAAGCGCTGCTGTG 990
Qy 343 Phe---GlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLys----- 356
Db 991 TTCAGTGGCGGCGCGCGCGCTCTCGACGTCGCGCGCAAGGAAATCCACATGGTC 1050
Qy 357 -----GlyAlaGlnAlaTyrAspGluTyrGlyArg 366
Db 1051 ATCCCGCGCGACCTCGCGGAGAACACGCTCAGGCAAGACGACGAGGTACGCG--- 1107
Qy 367 AspAspTyrSerSerArgThrLysAsnGlySerGlyAlaAspLys----- 382
Db 1108 -----GCAGTGGCATTGGTGGCGCGCGCGGAGAGACTTCAGCTTCGGA 1155
Qy 383 GlyGlyProThr----- 386
Db 1156 GCGCGCAAGACGGTGGACGCGCGCGGAGGAGTAGACGAGGCGCGCTTGCCTGACGGG 1215
Qy 387 LeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLys-----AspAsp 403
Db 1216 CTGACGAGATGGGTTCGAGCTCGACGCGGAGCTGCACCCGAGGTGCTGCGACTCGAC 1275
Qy 404 GlyGluGlyArgAlaAlaValAla-----MetProProAlaSer 417
Db 1276 GGAACCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1335
Qy 418 ValMetThrArgLeuIleLeuMetValTyrArgLysLeuIleArgAsnProAsnThr 437

```

```

Db 1336 GTGATGACGCGCTCATCTCTCATTAATGGTGTGGCGCAAGCTCATCCGCAACCAACT 1395
Qy 438 TyrSerSerLeuIleGlyValValTyrSerLeuValSerTyrArgTyrGlyIleGluMet 457
Db 1396 TACTCCAGCTCTCGGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCTCGCT 1455
Qy 458 ProAlaIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMet 477
Db 1456 CCAGCAATCGTCAGAGTCCATCTCTCTCGGACGAGCGCTGGGGATGGCCATG 1515
Qy 478 PheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleAlaCysGlyAsnLysLeu 497
Db 1516 TTTAGCTGGATGTTCATGGCGCTGCAGCCGAGCATCATCGCGTGGCAATCAGCC 1575
Qy 498 AlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSer 517
Db 1576 GCGCTGCTCTCCATGCGCGCTTCCTCGCGGCGCTCGCTCGCTCGCTCGCTCGCTCG 1635
Qy 518 IleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaLeuPro 537
Db 1636 ATGCCATCGATTCGCGGAGCTCTCGACAGTGCCTATGTTGAGGCGGTCTACCA 1695
Qy 538 GlnGlyIleValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSer 557
Db 1696 CAAGGATTCGCTTTTGTGTTTGCAGAAAGATAACAATGCCACCGCGCATCTGAGC 1755
Qy 558 ThrAla-----TyrGly-----ProIleThr 564
Db 1756 ACAGCGGTAATTTTGGCATGCTAATAGCTCTTCCCAATCACA 1797

RESULT 15
LOCUS AY093960 1869 bp mRNA linear PLN 27-APR-2002
DEFINITION Arabidopsis thaliana Atlg73590/F6D5_2 mRNA, complete cds.
ACCESSION AY093960
VERSION AY093960.1 GI:20334719
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 1869)
Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J.,
Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
Theologis,A. and Becker,J.R.
Arabidopsis ORF clones
Unpublished
2 (bases 1 to 1869)
Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J.,
Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D.,
Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G.,
Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M.,
Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T.,
Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C.,
Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W.,
Theologis,A. and Becker,J.R.
Direct Submission
Submitted (02-APR-2002) Salk Institute Genomic Analysis Laboratory
(SIGNAL), Plant Biology Laboratory, The Salk Institute for
Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
USA
Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.

```

The Salk, Stanford, FGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Cheuk, R., Chen, H., Kim, C. J., Meyers, M. C., Shinn, P., Barh, J., Bowser, L., Chan, M. M., Chang, E., Dale, J. M., Deng, J. M., Goldsmith, A. D., Jones, T., Karlin-Neumann, G., Lam, B., Lee, J. M., Lin, J., Miranda, M., Nguyen, M., Onodera, C. S., Palm, C. J., Quach, H. L., Southwick, A., Tang, C. C., Toriumi, M., Wu, H. C., Yamada, K., Yamamura, Y., Yu, G., Davis, R. W., Theologis, A., and Ecker, J. R.

Cheuk, R. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Ecker, J. R. (SSP/Salk) contributed equally to this work as PIs.

FEATURES  
source Location/Qualifiers  
1. .1869  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/chromosomes="1"  
/clone="U12338"  
/ecotype="Columbia"  
/notes="This clone is in pUNI 51"  
CDS  
1. .1869  
/notes="auxin transporter splice variant b, putative"  
/codon\_start=1  
/product="At1g73590/F6D5.2"  
/protein\_id="AAM16221.1"  
/db\_xref="GI:20334720"  
/translation="MITRADFTVMTVMPLVYMLAYGVKWKIETPQCQGINR  
FVALPAPLLSFHTAANPYANFLAASLQKVI VLSLLFLWCKLSRNSLDWTI  
TFLSLTPTNLYVGLPLKGMYPNSGLMVI VVLCIIVYTLMLFLFYRGAKLL  
ISGQPDTAGSVISIVHSDIMSLDGRQPLEATEIKEDKLHVTVRNSASRSDIYS  
RRSQGLSATPRSNLTNABISYLSQSRNPTPRGSGFNHTDFYSWASGGGRNSNFGPG  
EAVFGSGKTPRPSNVEEDGGPAKTAAGTAAGREHYQSGSGGGGGAHPVAPNPG  
MPSNPTGGGGTAAGKNA PVVGGKQDNGRDLHFMVSSSASPVSVDVFGGGGNHHA  
DYSTATNDHQVKVLSIPGNSNDQYVEREFPNGKDDSDSKVLATDGGNNISKTT  
QAKMPTSVMTLLILIMVRLIRNPNSISLFGITWLSIFKNWIEPALIAKSIYS  
ILSDAGLGMFSLGIFALNPRIIACGNRRFAAFAMRFVVGPAVMLVASYAVGLRG  
VLLHVAIIQAALPQGI PVFVFAKEYNVHFDILSTAVIFGMLIPLITLLYLLGL"

ORIGIN

Alignment Scores:  
Pred. No.: 3,96e-84 Length: 1869  
Score: 1647.50 Matches: 366  
Percent Similarity: 64.82% Conservative: 56  
Best Local Similarity: 56.22% Mismatches: 104  
Query Match: 56.62% Indels: 125  
DB: 8 Gaps: 18  
US-10-030-884-14 (1-573) x AY093960 (1-1869)

Qy 1 MetIleThrAlaLeuAspLeuTyHisValLeuThrAlaValValProLeuTyValAla 20  
Db 1 ATGATTACGGCGCGGACTTCTACCGCTTATGACGGCTATGTTCCGTATACGTAGCT 60  
Qy 21 MetThrLeuAlaTyGlySerValArgTptTptArgIlePheThrProAspGlnCysser 40  
Db 61 ATGATCTCTCGCTTACGGCTCTGTCATATGTGTGGAATCTTACACACCAATGCTCC 120  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 121 GGCAATAACCGTTTCGTCTCTCTTCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnTyVal 80  
Db 181 GCTAACACCGCTTACGGCATGACCTCCGTTCTCTCGCCGACGATTCTCTCAGAAAGTC 240  
Qy 81 AlaValLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaLeuGly 100  
Db 241 ATTGCT 291  
Qy 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
Db 292 TTAGATTGGACCAATACT 351

Qy 121 IleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
Db 352 ATACCTCTTCTCAAGGCGATGTATGGT---AATTCTCCGGCAGCTCATGTGTTCAATC 408  
Qy 141 ValValLeuGlnCysserIleTptTptThrLeuMetLeuPheLeuPheGluTyArgAla 160  
Db 409 GTTGTCTTTCAGTGATATCAITTTGGTACACTCATGCTCTTCTCTTGTGATACCGTGA 468  
Qy 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
Db 469 GCTAAGCTTTTGTATCTCCGAGCAGTTTCCAGAC---ACACGAGGATCTATTGTTGAT 525  
Qy 181 ArgValAspSerAspValValSerLeu---AlaArgGlyAspValGluLeuGluAlaGlu 199  
Db 526 CATGTTGATTCGACATTATGCTTTAGTAGGAGACAACTTTGGAAACGTAAGCTGAG 585  
Qy 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAlaGlyArgVal 219  
Db 586 ATTAAA-----GAAGATGGAGAGCTT 606  
Qy 220 ArgValThrValArgLySerThrSerArgSerGluAlaAlaCysSerHisSerHis 239  
Db 607 CATGTTTACTGTCGTCGTTCTTAATGCTTCAAGGTCTGATATTTACTCGAAGAGCTCAA 666  
Qy 240 SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTySerLeuGln 259  
Db 667 GGCTTATCTCGACACCTAGACCTTCGAATCTAACCAACGCTGAGATATATTCGCTTCA 726  
Qy 260 SerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIle 279  
Db 727 AGTTCAAGAAACCCACGCCACGCTGCTAGTTTTAAATCATATCATGATTTTACTCGATG 786  
Qy 280 Val----- 280  
Db 787 ATGCTCTCTGCTGCTGCTGGAACCTTAACTTTGCTCCTCGAGAAGCTGTTGTTGTTCT 846  
Qy 281 -----GlyAlaAlaAlaLys----- 285  
Db 847 AAGGTCCTACTCCGAGACCTTCCAACTACGAAGAAGACGGTGGTCTGCTGTAACCCGAG 906  
Qy 286 GlyGlyGlyGlyAlaAlaGly-----AspGluGluTyGlyAlaCysGlyGly 301  
Db 907 GCTGCTGGAACCTGCTGCTGAGCTGGAGGTTTCAATATCAATCTGGAAGTGGTGGC 966  
Qy 302 GlyGlyGlyGlyHisSerProGlnPro----- 310  
Db 967 GTTGAGGAGCGCATTTATCCGCGCGCCGAACCCAGGGATGTTTTCGCCCAACACTGGCGGT 1026  
Qy 311 -----GlnAlaValAlaValProAlaLysArg----- 319  
Db 1027 GGTGAGGCGCGCGCGGGAAGAAAGCTCCCGTGGTGGTGGGAAAGACAGACGGA 1086  
Qy 320 -----LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArg 337  
Db 1087 AACGGAAGAGATCTTTCATGTTTGTGGAGCTCAAGTGTTCGCGGCTCTCAGAT--- 1143  
Qy 338 AlaAlaValHisValPheGlyAlaGlyAlaAlaAspHisAlaAspValLeuAlaLysGly 357  
Db 1144 -----GTGTTCCGCGGTGGAGAGGAACCCACAC-----GCCGATTAC 1182  
Qy 358 AlaGlnAlaTyAspGluTyGlyArgAsp----- 367  
Db 1183 TCCACCGCTACGAACGATCATCAAAAGACGTTAAGATCTCTGTACCTCAGGGGAATAGT 1242  
Qy 368 -----AspTySerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGly 384  
Db 1243 AACGACAAACGATGCTGGAGAGGAGAGTTTGTGTTTCGTTAACAAGACGATGATAGC 1302  
Qy 385 ProThrLeuSerLysLeuGly-----SerAsnSerThrAlaGlnLeuTyPro 400  
Db 1303 AAGTATTGGCAACGCGCGTGGGAAACACATATGACCAACCAACGACGCGAG----- 1353

```
Qy 401 LysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThr 420
Db 1354 -----GCTAAGGTGATGCCACCACCAAGTGTGATGACA 1386
Qy 421 ArgLeuIleLeuIleMetValTirArgLysLeuIleArgAsnProAsnThrTyrSerSer 440
Db 1387 AGACTCATTCTCATATGGTTGGAGAACTTATTCGTAAATCCCACTCTTACTCCAGT 1446
Qy 441 LeuIleGlyValValTirpSerLeuValSerTyrArgTirpGlyIleGluMetProAlaIle 460
Db 1447 TTATTGGCATCACCTGCTCCTCATTTCTTCAAGTGAACATTGAAATGCCAGTCTTT 1506
Qy 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480
Db 1507 ATAGCAAGTCTATCTCCATCTCTCAGATGCAGGTTAGCATGGCTATGTTCACTCTT 1566
Qy 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaAlaIle 500
Db 1567 GGGTTGTTTCATGGCGTTAAACCCCAAGAAATAATAGCTTGTGAAACACAGAGCAGCTTTT 1626
Qy 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaVal 520
Db 1627 GCGCGGCTATGAGATTGTGTTGGACCTGCGCTCATGCTCGTGTCTTCTTATGCCGTT 1686
Qy 521 GlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIle 540
Db 1687 GGCCTCCGTGGCGTCTCTCCATGTTGCCATTATCCAGGCAGCTTGGCCGAAGGAATA 1746
Qy 541 ValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAla--- 559
Db 1747 GTACCGTTTGTGTTTGCCAAAGAGATATAATGTGCATCTCGACATTCTTAGCACTGCGGTG 1806
Qy 560 ---TyrGly-----ProIleThr 564
Db 1807 ATATTGGGATGTTGATCGCGTTCGCCATAACT 1839
```

Search completed: March 3, 2004, 11:11:39  
Job time : 6430 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 3, 2004, 09:12:55 ; Search time 3990 Seconds  
(without alignments)  
4288.479 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910

Sequence: 1 MITALDIYHLTVAVPLYVA.....DILSTAYGPTSGFTCHS 573

Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xih  
-Q=/cgn2\_1/USPTO.spool/US10030884/runat\_24022004\_103923\_1960/app.query.fasta\_1.711  
-DB=EST -QFWT=fastcap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10030884 @cgn\_1\_1906 @runat\_24022004\_103923\_1960 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:

1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hcti:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hcti:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_eston:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_phg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1922	66.0	2737	11 AY110494	AY110494 Zea mays
2	1405	48.3	840	29 CG358070	CG358070 OGBK16TV
3	1356	46.6	893	29 CG376782	CG376782 OGBK16TV
4	1222	42.0	835	29 CC722061	CC722061 OGUZ102TV
5	1204	41.4	730	29 CG376770	CG376770 OGUZ102TV
6	1100	37.8	776	29 CG317974	CG317974 OGWAG1TV
7	1043.5	35.9	658	13 BU098540	BU098540 946136B05
8	1008.5	34.7	627	13 BU037419	BU037419 946138P12
9	1004	34.5	1426	11 AY106004	AY106004 Zea mays
10	999.5	34.3	610	13 BU080190	BU080190 946150G03
11	997.5	34.3	604	14 CA831110	CA831110 1117015F0
12	983	33.8	630	14 CA830783	CA830783 1117011D1
13	962	33.1	726	29 CG358058	CG358058 OGBK16TH
14	936	32.2	1221	11 AY110024	AY110024 Zea mays
15	909	31.2	972	28 CG371166	CG371166 PUSK16TB
16	907	31.2	717	14 CF244269	CF244269 3530_1_28
17	885.5	30.4	903	14 CF243672	CF243672 3530_1_23
18	881.5	30.3	552	13 BU037594	BU037594 946140E03
19	848.5	29.2	1105	14 CK208792	CK208792 FGAS02051
20	837	28.8	1107	14 CK209475	CK209475 FGAS02124
21	836.5	28.7	1109	14 CK208849	CK208849 FGAS02057
22	819.5	28.2	739	12 EG600488	EG600488 EST506383
23	809.5	27.8	868	29 CG317965	CG317965 OGWAG1TH
24	805.5	27.7	497	13 BU036965	BU036965 946131B07
25	800.5	27.5	760	29 CC722052	CC722052 OGUZ102TV
26	798.5	27.4	728	14 CB348868	CB348868 CAB2SG000
27	791.5	27.2	796	9 AF162834	AF162834 AF162834
28	778.5	26.8	774	14 CF437477	CF437477 EST673822
29	769	26.4	963	29 CG861094	CG861094 ZMWB026
30	762	26.2	619	14 CA236508	CA236508 SCSEB1502
31	758	26.0	819	14 CB648277	CB648277 OSUNEB11J
32	752	25.8	741	13 BQ999371	BQ999371 OGUZ1008
33	748	25.7	861	14 CK279869	CK279869 EST725947
34	748	25.7	897	14 CK282170	CK282170 EST744892
35	745.5	25.6	931	14 CK293778	CK293778 EST756492
36	733.5	25.2	851	28 BH553421	BH553421 BOGDR35TF
37	731.5	25.1	702	28 BZ011975	BZ011975 Oe386B11
38	730	25.1	469	14 CD001628	CD001628 3529_1_10
39	729	25.1	696	14 CA184766	CA184766 SCSEB1309
40	727.5	25.0	633	14 CF805982	CF805982 psHB002x1
41	727.5	25.0	929	14 CK294982	CK294982 EST757696
42	724.5	24.9	739	14 CB646582	CB646582 OSUNEB09C
43	720.5	24.8	913	14 CK250728	CK250728 EST734365
44	718.5	24.7	901	14 CK292860	CK292860 EST755574
45	717.5	24.7	894	14 CK284513	CK284513 EST747235

# ALIGNMENTS

RESULT 1  
AY110494  
LOCUS Zea mays CL464\_1 mRNA sequence.  
DEFINITION Zea mays CL464\_1 mRNA sequence.  
ACCESSION AY110494  
VERSION AY110494.1 GI:21214903  
KEYWORDS HTC  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 2737)

**AUTHORS** Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S.,  
 Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.  
**TITLE** Maize Mapping Project/DuPont Consensus Sequences for Design of  
 Overgo Probes  
**JOURNAL** Unpublished (2002)  
**REFERENCE** 2 (bases 1 to 2737)  
**AUTHORS** Coe,E.H.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (25-APR-2002) Maize Mapping Project, University of  
 Columbia, MO 65211, USA  
**COMMENT** If you are interested in getting corresponding physical clones,  
 these are publicly available from ZmDB, [www.zmdb.iastate.edu](http://www.zmdb.iastate.edu); TIGR,  
[www.tigr.org](http://www.tigr.org); or NCBI, [www.ncbi.nlm.nih.gov](http://www.ncbi.nlm.nih.gov). When the source of the  
 maize cDNA sequences is either Virginia Walbot, Stanford or Pat  
 Schnable, Iowa State, then clones may be requested from ZmDB:  
[www.zmdb.iastate.edu](http://www.zmdb.iastate.edu).

**FEATURES**  
 source  
 1..2737  
 Location/Qualifiers  
 /organism="Zea mays"  
 /mol\_type="mRNA"  
 /db\_xref="MaizeDB:632106"  
 /db\_xref="taxon:4577"  
 /clone\_lib="Maize Mapping Project/DuPont Consensus  
 Library"  
 /note="this sequence is part of a project of EST  
 assemblies resulting from the application of public  
 contigs to seed Dupont contigs; this resource was  
 assembled by Dupont as part of a collaboration for the  
 overgo addressing of BACs in conjunction with the Maize  
 Mapping Project"

## ORIGIN

Alignment Scores:  
 Pred. No.: 4,75e-135 Length: 2737  
 Score: 1922.00 Matches: 406  
 Percent Similarity: 70.98% Conservative: 0  
 Best Local Similarity: 70.98% Mismatches: 166  
 Query Match: 66.05% Indels: 0  
 DB: 11 Gaps: 0

US-10-030-884-14 (1-573) x AY110494 (1-2737)

QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 DB 171 ATGATCACCGCGTGGACCTCTACACAGTGTGTGACGGGGTGGTGGTGTGACGTGCC 230  
 QY 21 MetThrLeuAlaTyrGlySerValArgTTPTrpArgIlePheThrProAspGlnCysSer 40  
 DB 231 ATGACGCTGGGTACGGCTCCGTCCGCTGGTGGCGCATCTTCACGCGGACGAGTGCTCC 290  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 DB 291 GGGATCAACCGCTTCGTGGCGCTCTCCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 350  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAspThrIleuGlnIysVal 80  
 DB 351 NCCAAAGACCCCTTCGCCATGACCTCGCTTCCTGGCCGCCGACACGCTGCAGAGGTG 410  
 QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 DB 411 GCCGTCTCGCGTGTGGCGCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 470  
 QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrIleuValMetGly 120  
 DB 471 NNNNNCTGGAGCATCAGCCTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 530  
 QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrIleuMetValGlnVal 140  
 DB 531 ATCCCGCTGTGGAGGATGATACGGCGGTGTGTGGCGCGGACGCTCATGGTCCAGNNN 590  
 QY 141 ValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160

DB 591 NNN 650  
 QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaSerIleValSerPhe 180  
 DB 651 NNN 710  
 QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
 DB 711 NNN 770  
 QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
 DB 771 GNN 830  
 QY 221 ValThrValArgIysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHisSer 240  
 DB 831 NNCACCGTCGCGAAGTCCACAGCTCGCGTCGAGGCGCGCTGCTGCACCTCGCACTCC 890  
 QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260  
 DB 891 CAGACCATGACGCCCGTGTGTCCAACTCTCCGGGTGGAGATCTACTCGCTGCAGTCG 950  
 QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIleVal 280  
 DB 951 NNN 1010  
 QY 281 GlyAlaAlaAlaIysGlyGlyGlyAlaAlaGlyAspGluGluIysGlyAlaCysGly 300  
 DB 1011 GGGCGCCGCCNN 1070  
 QY 301 GlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAlaIysArgIys 320  
 DB 1071 NNN 1130  
 QY 321 AspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAlaVal 340  
 DB 1131 GACCTGCACATGCTCGTCTGGAGCTCAGCGCTCCGCCGTGTCGAGGCGCGCGCGTGG 1190  
 QY 341 HisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaIysGlyAlaGlnAla 360  
 DB 1191 CACGTCTTCGCGCCGCGCGCTGACCATGCGAGCTCTCGCCAAAGAGAGCCAGGCC 1250  
 QY 361 TyrAspGlnTyrGlyArgAspAspTyrSerSerArgThrIysAsnGlySerGlyAla 380  
 DB 1251 TACGACGAGTACGGCGCGCAGCTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1310  
 QY 381 AsplysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrPro 400  
 DB 1311 NNN 1370  
 QY 401 LysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProAlaSerValMetThr 420  
 DB 1371 NNN 1430  
 QY 421 ArgLeuIleLeuIleMetValTTPArgLysLeuIleArgAsnProAsnThrTyrSerSer 440  
 DB 1431 CGGCTCATCTCATCTCATCTGTGTGAGAGAGCTGATCCGGAAACCCCAACACCTACTCC 1490  
 QY 441 LeuIleGlyValValTrpSerLeuValSerTyrArgTTPGlyIleGluMetProAlaIle 460  
 DB 1491 CTATCGCGCTCGTCTGGTCCCTGGTCTCTA CAGGTGGGCGCATCGAGATGCCAGGATC 1550  
 QY 461 IleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeu 480  
 DB 1551 ATCGCCCGGTCGATTTGATCTCTGTGGACCGGGTCTCGGGATGGCCATGTTACGCCCTA 1610  
 QY 481 GlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIle 500  
 DB 1611 GGCGTGTTCATGCGCTGCAGCCGAGGATCATCGCTGGGGAAACAGCTGGCGGCGCATC 1670  
 QY 501 AlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaIleSerIleAlaVal 520  
 DB 1671 GCGATGGGCGTCCGGTTCGTGCGAGGCCCGCGGTTCATGGCCCGCGCTTCATCGCGCTC 1730

```

Qy 521 GlyLeuArgGlyValLeuLeuHisLeAlaIleValGlnAlaAlaLeuProGlnGlyIle 540
Db 1731 GGCTCGCGCGGCTCCTCCATCCATCGCCATCGTCCAGGCTGCTGCTCAGGGGATC 1790
Qy 541 ValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAlaTyr 560
Db 1791 GTCCCGTTCGTGTCGCAAGAGTACGGGGTTCATCCCGACATCCTGACACAGCGTAT 1850
Qy 561 GlyProIleThrSerHisGlyPheIleThrCysHis 572
Db 1851 GGTCCATATACATCGCATGGTTTCATCATTCGTCAT 1886

RESULT 2
CG358070 940 bp DNA linear GSS 26-AUG-2003
LOCUS OGI1K16TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0724C07,
DEFINITION genomic survey sequence.
ACCESSION CG358070
VERSION 1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 840)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OGI1K16TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..840
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone_lib="ZMMBma0724C07"
/clone_lib_ZM_0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.: 1,53e-96 Length: 840
Score: 1405.00 Matches: 279
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.28% Indels: 0
DBs: 29 Gaps: 0

US-10-030-884-14 (1-573) x CG358070 (1-840)

Qy 74 AlaAspThrLeuGlnLysValAlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeu 93
Db 838 GCCGACACGCTGCAGAGGTGCCGCTCTCGCGTCTGCGGCTGGCGCTCCCGCGGCTC 779
Qy 94 SerSerProArgAlaLeuGlyLeuAspThrSerIleThrLeuPheSerLeuSerThrLeu 113
Db 778 TCCCTCCCGCGCGCTCGGGCTCGACTCGAGCATCAGCTCTCTCTCCCTCCACGCTC 719
Qy 114 ProAsnThrLeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerAla 133

```

```

Db 718 CCCAACCGCTCTGTCATGGGCATCCCGCTGTGCGAGGCATGTACGGCGCGCTGTCGGCC 659
Qy 134 GlyThrLeuMetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeu 153
Db 658 GGCACGCTCATGGTCCAGTCTGCTCTCCAGTGCATCATCTGTCACACGCTCATGCTC 599
Qy 154 PheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAla 173
Db 538 TTCTCTTTTCAGTACCGCGCGCGCGCTGCTGCTGCTGACCACTCTCCCGACGGCGCC 539
Qy 174 AlaAlaSerIleValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAsp 193
Db 538 GCGCGTCCATCTCTCTCTCCGCTGACTCCGAGTCTGCTGCTGCGCAGGGGGAC 479
Qy 194 ValGluLeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGly 213
Db 478 GTGAGCTCGAGCGCGCGCGCGCTGCGCGCGCGCGCGCTCTCTCCCGCGCGC 419
Qy 214 GlyAspAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAla 233
Db 418 GGGGACGCGCGCGGCTGCGGTCACTCCGCTGCGAAGTCCACGCTCGCTGCCAGGCC 359
Qy 234 AlaCysSerHisSerHisSerGlnThrMetGlnProArgValSerAsnLeuSerGlyVal 253
Db 358 GCGTCTCGCACTCCGCACTCCGACCATGAGCCCGCGCTGTGTCCAACTCTCCCGCGCTG 299
Qy 254 GluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHis 273
Db 298 GAGATCTACTCGTCTCAGTCTGCGCGCAACCCCGCGCGCTGCTCCAGCTTCAACAC 239
Qy 274 AlaAspPheAsnIleValGlyAlaAlaLysGlyGlyGlyAlaAlaGlyAsp 293
Db 238 GCCGACTTCTTCAACATCTGCGCGCGCGCGCAAGGGAGCGGAGGCGCGCGCGGAC 179
Qy 294 GluGluLysGlyAlaCysGlyGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaVal 313
Db 178 GAGGAGAGAGCGCGCATGCGCGCGCGCGCGCGGAGAGACACTCGCCGACGCGCGCGCTC 119
Qy 314 AlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerSerAlaSerPro 333
Db 118 GCGGTGCGCGCCACAGAGGAGACCTGCACATGCTGCTGAGAGTCCAGCGCTCGCGCC 59
Qy 334 ValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAsp 352
Db 58 GTGTCGAGCGCGCGCGCTGCACGTCTTGGCGCGCGCGCGCTGACCATGCGCGAC 2

RESULT 3
CG376782 893 bp DNA linear GSS 26-AUG-2003
LOCUS OG3CN22TV ZM_0.7_1.5_KB Zea mays genomic clone ZMMBma0776C20,
DEFINITION genomic survey sequence.
ACCESSION CG376782
VERSION 1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 893)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T.,
Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N.
Consortium for Maize Genomics
Unpublished (2002)
Other_GSSs: OG3CN22TH
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF

```

Class: sheared ends.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..893  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone\_lib="ZM0776C20"  
 /clone\_lib="ZM 0.7 1.5 KB"  
 /note="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

## ORIGIN

Alignment Scores:  
 Pred. No.: 8,84e-93 Length: 893  
 Score: 1356.00 Matches: 273  
 Percent Similarity: 99.64% Conservatives: 1  
 Best Local Similarity: 99.27% Mismatches: 0  
 Query Match: 46.60% Indels: 1  
 DB: 29 Gaps: 0

US-10-030-884-14 (1-573) x CG376782 (1-893)

QY 1 MetTleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 DB 69 ATGATCACCGCGCTGGACCTTACACAGTGTGACGGGGTGTGCTCCCTGTAGTGCC 128  
 QY 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40  
 DB 129 ATGACGCTGGCTACGGCTCCGTCGCTGGTGGCGCATCTTCAACGCGGACGAGTGCTCC 188  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 DB 189 GGGATCAACCGCTTCGTGGCGCTCTTCGCGGCTGCTCCGCGGCTCTCTCCCTTCATCTTC 248  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 DB 249 ACCAAGACCCCTTGGCATGAACTTGGCTTCTTGGCGCGGACACGCTCGACAGGTG 308  
 QY 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 DB 309 GCGGCTCTCGCGCTGTGGCGTGGCTCCCGCGGCTCTCTCCCGCGCGCGCTCGG 368  
 QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
 DB 369 CTCGACTGGAGCATACGGCTCTTCCCTCTCCAGCTTCCCAACACGCTCGTCATGGGC 428  
 QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 DB 429 ATCCGCTCTGCGAGGCATGTACGCGGCTGTGCGCGGCGGACGCTCATGTCCAGGTC 488  
 QY 141 ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
 DB 489 GTCGCTCTCCAGTGCATCTGGTACACGCTCATGCTCTTCTCTTCGAGTACCGGCC 548  
 QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 DB 549 GCGCGCGCGCTCTGCTCCGACCATGTTCCCGGCGCGCGCGCTCCATGCTCTCTTC 608  
 QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200  
 DB 609 CGCGTCGACTCGACGCTGTCTGCTGCGCAGGGGGAGCTGAGCTCGAGGCCGAGGCC 668  
 QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220  
 DB 669 GACGGGTCGCGCGCGCGCGCTCTCTCCCGCGGGGACCGCGGGGGGTGCGC 728  
 QY 221 ValThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSerHisSer 240  
 DB 729 GTACCGTGCAGTCCAGTCCACAGCTCGCGTCCGAGGCGCGGTGCTCGACATCGACATCC 788  
 QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260

DB 789 CAGACCATGCGAGCCCGGTGTGTCCAACTCTCGGCGTGGAGATCTACTCGTGCAGTCG 848  
 QY 261 SerArgAsnProThr-ProArgGlySerSerPheAsnHisAla 274  
 DB 849 TCGGCAGACCCACCCCGCGCGGTCCAGCTTTCACACGCC 891

RESULT 4  
 CC722061 835 bp DNA linear GSS 19-JUN-2003  
 LOCUS OGUIZ09TV ZM 0.7 1.5 KB Zea mays genomic clone ZM0452A18,  
 DEFINITION genomic survey sequence.  
 ACCESSION CC722061  
 VERSION CC722061.1 GI:32126837  
 KEYWORDS GSS.  
 SOURCE Zea mays  
 ORGANISM Zea mays  
 BUKARYOTA; VIRIDIPANTAE; STREPTOPHYTA; EMBRYOPHYTA; TRACHEOPHYTA;  
 SPERMATOPHYTA; MAGNOLIOPHYTA; LILIOPSIDA; POALES; POACEAE; PACCAD  
 CLADE; PANICOLIDEAE; ANDROPOGONEAE; ZEA.  
 REFERENCE 1 (bases 1 to 835)  
 White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
 Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,  
 Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.  
 Consortium for Maize Genomics  
 Other\_GSSs: OGUIZ09TH  
 Unpublished (2002)  
 Contact: Cathy Whitelaw

TIGR  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301-838-5843  
 Fax: 301-838-0208  
 Email: whitelaw@tigr.org  
 Seq primer: TP  
 Class: sheared ends.  
 Location/Qualifiers  
 1..835  
 /organism="Zea mays"  
 /mol\_type="genomic DNA"  
 /strain="B73"  
 /db\_xref="taxon:4577"  
 /clone\_lib="ZM0452A18"  
 /clone\_lib="ZM 0.7 1.5 KB"  
 /note="Vector: pBCSK-; Site 1: HindII; 0.7-1.5 kb  
 methylation filtered genomic DNA library"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,21e-82 Length: 835  
 Score: 1222.00 Matches: 245  
 Percent Similarity: 88.13% Conservatives: 0  
 Best Local Similarity: 88.13% Mismatches: 0  
 Query Match: 41.99% Indels: 33  
 DB: 29 Gaps: 1

US-10-030-884-14 (1-573) x CC722061 (1-835)

QY 145 CysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAlaAlaArgAlaLeu 164  
 DB 2 TGCATCATCTGGTACACGCTCATGCTCTTCTTCGAGTACCGCGCGCGCGGCTC 61  
 QY 165 ValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPheArgValAspSer 184  
 DB 62 GTCCTGACCAAGTTCCTCCGACGCGCGCGCGCTCCATCGTCTCTTCCGCGTGCATCC 121  
 QY 185 AspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluProAspGlyValAla 204  
 DB 122 GACGTCGCTCGTTCGTCGCGAGGGGGAGCTGAGCTCGAGCGCGCGCGCGGCTGCC 181  
 QY 205 GlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArgValArg 224  
 DB 182 GCGCGCGCGCGCTCTCTCCCGCGGGGAGCGCGCGCGGTGCGGTGCGGTGCGGTGCGG 241  
 QY 225 LysSerThrSerArgSerGluAlaAlaCysSerHisSerHisSerGlnThrMetGln 244

```

242 AAGTCCACAGCTCGCGCTCGAGGCGCGTCTCGCATCTCGCACTCCAGACCATGAG 301
245 ProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnPro 264
302 CCGCGTGTCTCAACCTCTCCGCGGTGGAGATCTACTCGCTGCAGTCGTCGCGCAACCC 361
265 ThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleValGlyAlaAla 284
362 ACCCGCGCGGGTCCAGCTTCAACACGCGGAGCTTCTTCAACATCGTGGCGCGCGCC 421
285 LysGlyGlyGlyGlyAlaAlaGlyAspGluGluLysGlyAlaCysGlyGlyGlyGly 304
422 AAGGAGCGCGAGGAGCGCGGGGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
305 GlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLysAspLeuHisMet 324
482 GGACACTCGCGCAGCGCAGCGCGCGCTCGCGCGCGCCCAAGAGGAGGAGGAGGAGGAG 541
325 LeuValTrpSerSerAlaSerProValSerGluArgAlaAlaValHisValPheGly 344
542 CTGCTGGAGCTCCAGCGCTCGCGCGGTGCGAGCGCGCGCGCGCGCGCGCGCGCGCG 601
345 AlaGlyGlyAlaAspHisAlaAspValLeuAla-Lys----- 356
602 GCCGGCGCGCTGACCATGCGGACGCTCTCGCCMAAGTTGGTAACATGCGTGATCTTT 661
356 ----- 356
662 TGCTTGATTTCTAACCAAAATGTTGCTGTGCAATTTTATTGTCAACCAATGGGTGTG 721
357 -----GlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerAr 372
722 CACAACACAAAGCAGGAGCGCGCGCTTACGACGAGTACGCGCGCGCAGGACTACAGCAG 781
372 gThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLys 389
782 GACGAAGAACGGAGCGCGCGCGCGGACAAAGCGCGCGCGCGCGCGCGCGCGCGCGCG 833

```

```

RESULT 5
CG376770/c 730 bp DNA linear GSS 26-AUG-2003
LOCUS CG3CN22TH ZM_0.7_1.5_KB Zea mays genomic clone ZMMBMA0776C20,
DEFINITION genomic survey sequence.
ACCESSION CG376770
VERSION CG376770.1 GI:34294037
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 730)
White, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D., and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: CG3CN22TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1. .730
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"

```

## FEATURES

source

```

/clone="ZMMBMA0776C20"
/clone_lib="ZM_0.7_1.5_KB"
/notes="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 2,3e-81 Length: 730
Score: 1204.00 Matches: 237
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 41.37% Indels: 0
DB: 29 Gaps: 0

```

US-10-030-884-14 (1-573) x CG376770 (1-730)

```

QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140
Db 729 ATCCCGCTCTCGAGGCGCATGTACGGCGGTGTCGGCGCGCACGCTCATGGTCCAGGTC 670
QY 141 ValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160
Db 669 GTGCTCTCCAGTGCATCATCTGGTACAGCTCATGCTCTCTCTCTCTCTCTCTCTCTCT 610
QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180
Db 609 GCGCGCGCGCTCTGCTCTGACAGTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 550
QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGluAlaGluPro 200
Db 549 CGCGTGCAGTCCGACGTCGCTCGCTCGCGAGGGGGGAGCGTCGAGCTCGAGCGCGAGCCC 490
QY 201 AspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgValArg 220
Db 489 GACGGCGTTCGCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 430
QY 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHisSer 240
Db 429 GTCCGCTGCGCAAGTCCACGAGCTCGCGCTCCGAGGCGCGCGCTCTCTCTCTCTCTCTCT 370
QY 241 GlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSer 260
Db 369 CAGACATGACGCGCGCGGTGTCCAACTCTCCGCGGTGGAGATCTACTCTGCTGCACTCG 310
QY 261 SerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleVal 280
Db 309 TCGCGCAACCCACCCCGCGCGGTTCAGCTTCAACACGCGCGACTTCTTCAACATCGTC 250
QY 281 GlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAspGluGlyGlyAlaCysGly 300
Db 249 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 190
QY 301 GlyGlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLys 320
Db 189 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 130
QY 321 AspLeuHisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaAlaVal 340
Db 129 GACCTGCACATGCTGCTGTGAGCTCCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 70
QY 341 HisValPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGly 357
Db 69 CACGCTTTCGGCGCGCGCGCGCTGACATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 19

```

## RESULT 6

CG317974

LOCUS

DEFINITION

CGMGA9ITV ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0570013,

genomic survey sequence.

ACCESSION

CG317974

VERSION

CG317974.1 GI:34235240

KEYWORDS

GSS.

SOURCE

Zea mays

CG317974 776 bp DNA linear GSS 26-AUG-2003  
CGMGA9ITV ZM\_0.7\_1.5\_KB Zea mays genomic clone ZMMBMA0570013,  
genomic survey sequence.

ACCESSION CG317974  
VERSION CG317974.1 GI:34235240  
KEYWORDS GSS.  
SOURCE Zea mays

```

ORGANISM      Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
REFERENCE     1 (bases 1 to 776)
AUTHORS      Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Reenick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.
CONSORTIUM   Consortium for Maize Genomics
JOURNAL       Unpublished (2002)
COMMENT      Other GSSs: OGWGA91TH
              Contact: Cathy Whitelaw
              TIGR
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-838-5843
              Fax: 301-838-0208
              Email: whitelaw@tigr.org
              Seq primer: TF
              Class: sheared ends.

FEATURES     source
             1..776
               /organism="Zea mays"
               /mol_type="genomic DNA"
               /strain="B73"
               /db_xref="taxon:4577"
               /clone="ZM8BMA0570013"
               /clone_lib="ZM 0.7 1.5 kb"
               /note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
               methylation filtered genomic DNA library"

ORIGIN
Alignment Scores:
Pred. No.:      2,016-73      Length:      776
Score:          1100.00      Matches:    225
Percent Similarity: 86.87%      Conservative: 0
Best Local Similarity: 86.87%      Mismatches: 1
Query Match:    37.80%      Indels:     34
DB:             29      Gaps:      1

US-10-030-884-14 (1-573) x CG317974 (1-776)

QY      209 ValSerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThrSer 228
Db      1 GTCTCTCCGCGCGGGACGCGCGGGCGGGTGGCGTCCAGTCCAGTCCACACG 60

QY      229 SerArgSerGluAlaAlaCysSerHisSerHisSerGlnThrMetGlnProArgValSer 248
Db      61 TCGCGCTCCGAGCGCGTGTCTCGCACTCGCACTCCAGACCATCGACCGCCGCTGTGCC 120

QY      249 AsnLeuSerGlyValGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 268
Db      121 AACCTCTCCGCGGTGGAGATCTACTCGTGCAGTGTGTGCGCAACCCACCCCGCGCGG 180

QY      269 SerSerPheAsnHisAlaAspPhePheAsnIleValGlyAlaAlaLysGlyGlyGly 288
Db      181 TCCAC-TTCAACCAACCGGACTTCTTCAACATCGTCCGCGCGCGCGCAAGGAGCGCGGA 239

QY      289 GlyAlaAlaCysAspGluGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 308
Db      240 GGAGCGCGCGGGACGAGAGAGAGGCGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 299

QY      309 GlnProGlnAlaValAlaValProAlaLysArgLysArgLysArgLysMetLeuValTrpSer 328
Db      300 CAGCGCGAGCGCGTCCGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 359

QY      329 SerSerAlaSerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAla 348
Db      360 TCCAGCGCGCTCGCGCGGTTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 419

QY      349 AspHisAlaAspValLeuAla-Lys----- 356
Db      420 GACCATGCCGACGCTCTCTCCGCCAAGGTTGGTAACAATGCGTGATCTTTTGTGTTGTTTC 479

```

```

QY      356 ----- 356
Db      480 CTACCAAAATGTTGTCTCTGCAATTTTATTGTACCAATGGTGTGCACACACAAGC 539

QY      357 -GlyAlaGlnAlaTyArgpGluTyArgpAspTyArgpSerArgpThrLysAsnG1 376
Db      540 AGGAGCCCGCGGCTACGACGAGTACGCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCG 599

QY      376 YSerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAl 396
Db      600 GAGCGCGCGCGGACAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 659

QY      396 aGlnLeuTyProLysAspAspGlyGlyArgAlaAlaAlaAlaAlaAlaMetProAla 416
Db      660 GCAGCTGTATCCCAAGGACGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 719

QY      416 aSerValMetThrArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 434
Db      720 GAGCGTGATGACCGGCTATCTCTCATCATCTGTTGAGGAGGAGTGTATCGGAC 774

RESULT 7
LOCUS     BU098540                658 bp     mRNA     linear     EST 29-AUG-2002
DEFINITION 946136B05.v1 946 - tassels primordium prepared by Schmidt lab Zea
           mays cDNA, mRNA sequence.
ACCESSION BU098540
VERSION    BU098540.1 GI:22546229
KEYWORDS   EST.
SOURCE     Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoidae; Andropogoneae; Zea.
            1 (bases 1 to 658)
REFERENCE   1 (bases 1 to 658)
AUTHORS     Walbot, V.
TITLE       Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL     Unpublished (1999)
COMMENT     Unpublished (1999)
            Contact: Walbot V
            Department of Biological Sciences
            Stanford University
            855 California Ave, Palo Alto, CA 94304, USA
            Tel: 650 723 2227
            Fax: 650 725 8221
            Email: walbot@stanford.edu
            Plate: 946136 row: B column: 05.
FEATURES   source
            1..658
              /organism="Zea mays"
              /mol_type="mRNA"
              /cultivar="OH43"
              /db_xref="taxon:4577"
              /tissue_type="tassels"
              /dev_stage="just after the transition from vegetative to
              inflorescence development"
              /lab_host="XLOLR"
              /clone_lib="946 - tassels primordium prepared by Schmidt
              lab"
              /note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;
              Site 2: XhoI; George Chuck dissected immature tassels
              between 1mm and 3mm. Sharon Stanfield prepared the cDNA
              library in HybriZAP. Sample insert size range was 350 bp
              to 3 kb with a 1 kb average."

ORIGIN
Alignment Scores:
Pred. No.:      3,08e-69      Length:      658
Score:          1043.50      Matches:    214
Percent Similarity: 98.17%      Conservative: 0
Best Local Similarity: 98.17%      Mismatches: 3
Query Match:    35.86%      Indels:     2
DB:             13      Gaps:      1

```

US-10-030-884-14 (1-573) x BU098540 (1-658)

```

Qy 323 HisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaValHisVal 342
Db 7 CACATGCTGCTGGAGCTC-AGGCGCTCCCGTCTCCGAGCGCGCGGTGACGTC 65
Qy 343 PheGlyAlaGlyAlaAspHisAlaAspValLeuAlaValGlyAlaGlnAlaTrpAsp 362
Db 66 TTGGCGCCCGCGCGCTGACCATGCGAGCTCTCTGCCAAGAGAGCCAGGCTACGAC 125
Qy 363 GluTrpGlyArgAspAspTrpSerSerArgThrLysAsnGlySerGlyAlaAspLys 382
Db 126 GAGTACGGCGCGACGACTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 185
Qy 383 GlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTrpLysAsp 402
Db 186 GGGGGCGCGAGCTGTCCAGCTGGGTCCAACTCGACGGCGCGAGCTGTACCCCAAGGAC 245
Qy 403 AspGlyGluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMetThrArgLeu 422
Db 246 GACGGCGAGGGAGGCGAGCGCGG---GCCATGCGCGCGCGAGCGGTGATGACGGCGTC 302
Qy 423 IleLeuIleMetValTrpA-gLysLeuIleArgAsnProAsnThrTrpSerSerLeuIle 442
Db 303 ATCTCATCATGCTGTGGAGAAAGCTGATCCGGAACCCCAACACTTACTCCAGCCTCATC 362
Qy 443 GlyValValTrpSerLeuValSerTrpArgTrpGlyIleGluMetProAlaIleIleAla 462
Db 363 GGGTGTCTGGTCCCTCGTATCTACAGGTGGGCGATCGAGATCGCGCATCGCGC 422
Qy 463 ArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeu 482
Db 423 CGGTGCAATTTCGATCTCTCGAGCGCGGTCTCGGATGGCCATGTTACGCTAGGCTG 482
Qy 483 PheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMet 502
Db 483 TTATGCGCTGCGACCGAGGATCATCGCTGCGGAAACGCTGGCGGCATCGCGATG 542
Qy 503 GlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeu 522
Db 543 GGGCTCGGTTCTGTCGAGCGCGCGGTCTATGGCGCGCGCTCCATCGCGCTGCTG 602
Qy 523 ArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIle 540
Db 603 CGCGGCTGCTCTCTCCATCGCATCGCATCGTCCAGGCTGCTCTGCTCAGGGATC 656

```

## RESULT 8

BU037419 946138F12.y1 946 - tassal primordium prepared by Schmidt lab Zea  
 DEFINITION 627 bp mRNA linear EST 23-AUG-2002  
 mays cDNA, mRNA sequence.

ACCESSION BU037419  
 VERSION BU037419.1 GI:22472939  
 KEYWORDS EST.

## SOURCE

Zeas mays  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
 clade; Panicoideae; Andropogoneae; Zea.  
 1 (bases 1 to 627)

## REFERENCE

Zeas mays  
 Maize ESTs from various cDNA libraries sequenced at Stanford  
 University

## JOURNAL

Unpublished (1999)  
 Contact: Walbot V  
 Department of Biological Sciences  
 Stanford University  
 855 California Ave, Palo Alto, CA 94304, USA  
 Tel: 650 723 2227  
 Fax: 650 725 8221  
 Email: walbot@stanford.edu

Plate: 946138 row: F column: 12.

Location/Qualifiers

## FEATURES

source 1..627

/organism="Zea mays"  
 /mol\_type="mRNA"  
 /cultiivar="OH43"  
 /db\_xref="taxon:4577"  
 /tissue\_type="tassels"  
 /dev\_stage="just after the transition from vegetative to  
 inflorescence development"  
 /lab\_host="XLOLR"  
 /clone\_lib="946 - tassal primordium prepared by Schmidt  
 lab"  
 /note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;  
 Site 2: XhoI; George Chuck dissected immature tassels  
 between 1mm and 3mm. Sharon Stanfield prepared the cDNA  
 library in HybriZAP. Sample insert size range was 350 bp  
 to 3 Kb with a 1 Kb average."

## ORIGIN

Alignment Scores:  
 Pred. No.: 1,31e-66 Length: 627  
 Score: 1008.50 Matches: 200  
 Percent Similarity: 97.10% Conservative: 1  
 Best Local Similarity: 96.82% Mismatches: 3  
 Query Match: 34.66% Indels: 3  
 DB: 13 Gaps: 2

US-10-030-884-14 (1-573) x BU037419 (1-627)

```

Qy 274 AlaAspPheAsnIleValGlyAlaAlaLysGlyGlyAlaAlaGlyAsp 293
Db 15 GCGGACTTCTTCAACATCGTCTTTCCGCCCCCAAGGAGGCGGAGCGCGGGAC 74
Qy 294 GluGluLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnAlaVal 313
Db 75 GAGGAGAGGGCGCATGC-----GGCGGCGAGGACACTCGCCGACCGCAGCCGTC 128
Qy 314 AlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPro 333
Db 129 GCGCTGCGCGGCAAGAGAGAGGACCTGCACATGCTCGTCTCGAGCTCCAGCGCTCGCCC 188
Qy 334 ValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAspVal 353
Db 189 GTCTCCGAGCGCGCGCGCTGACAGTCTTCGCGCGCGCGCGCTGACCATCGCGACGTC 248
Qy 354 LeuAlaLysGlyAlaGlnAlaTyAspGluTyArgAspAspTySerSerArgThr 373
Db 249 CTCGCCAAGAGAGCGCCAGCCCTACGACGATACGCGCGCGCGCGCTACAGCAGGAGC 308
Qy 374 LysAsnGlySerGlyGlyAlaAspLysGlyProThrLeuSerLysLeuGlySerAsn 393
Db 309 AAGAAGGAGCGCGCGCGCGCGAGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 368
Qy 394 SerThrAlaGlnLeuTyProLysAspAspGlyGluGlyArgAlaAlaAlaValAlaMet 413
Db 369 TCGACGGCGCGAGCTGTACCCCAAGAGCAGCGCGCGAGGAGGAGCGCGCGCGCG 425
Qy 414 ProProAlaSerValMetThrArgLeuIleLeuIleMetValTrpArgLysLeuIleArg 433
Db 426 CGCGCGCGGAGCGGTGATGACCGGCTCATCTCATCATGTGTGTGAGGAAGCTGATCCGG 485
Qy 434 AsnProAsnThrTySerSerLeuIleGlyValValTrpSerLeuValSerTyArgTrp 453
Db 486 AACCCCAACACATTACTCCAGGCTCATCGCGTCTGCTGCTGCTGCTGCTGCTGCTGCT 545
Qy 454 GlyIleGluMetProAlaIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeu 473
Db 546 GGCATCGAGATGCCAGCGATCATGCCCGGCTGATTTTCATTCCTGCTCGGAGCGGGTCTC 605
Qy 474 GlyMetAlaMetPheSerLeu 480
Db 606 GGGATGGCCATGTTTCAGCCTA 626

```

RESULT 9

AY106004	LOCUS	AY106004	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
DEFINITION		AY106004	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
ACCESSION		AY106004	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
VERSION		AY106004.1	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
KEYWORDS		HTC	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
SOURCE		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
ORGANISM		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
REFERENCE		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
AUTHORS		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
TITLE		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
JOURNAL		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
REFERENCE		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
AUTHORS		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
TITLE		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
JOURNAL		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
COMMENT		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
FEATURES		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
source		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
ORIGIN		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Alignment Scores:		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Pred. No.:		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Score:		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Percent Similarity:		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Best Local Similarity:		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Query Match:		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
DB:		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
US-10-030-884-14 (1-573) x AY106004 (1-1426)		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
QY		Zea mays	Zea mays	1426 bp	mRNA	linear	HTC 16-OCT-2002
Db							

D	b		420	GCGCTGCTTGGTCCTGGATCCTACAGGTGGGCATCGAGATCCAGCGATCATGCC	479
Q	y		463	ArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetalaMetPheSerLeuGlyLeu	482
D	b		480	CGETCGATTTCGATCTCTGTCGGACCGGGTCTCGGATGACCATGTTTCAGCCTAGGCGTG	539
Q	y		483	PheMetaLeuGlnProArgIleIleAlaCySgLyVasnLysLeuAlaAlaIleAlaMet	502
D	b		540	TTCATGCGCTGCAGCGGAGATCATCGCTGCGGAACAAGTCGGCGGCATCGCGATG	599
Q	y		503	Gly 503	
D	b		600	GGG 602	
 RESULT 11					
CA831110                      604 bp    mRNA       linear     EST 12-DEC-2002					
LOCUS                      1117015F07.y1 1117 - UniGene V from Maize Genome Project Zea mays					
DEFINITION                  cDNA, mRNA sequence.					
ACCESSION                   CA831110					
VERSION                    1 GI:26558875					
KEYWORDS                   EST.					
SOURCE                    Zea mays					
ORGANISM                   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
REFERENCE                   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD					
AUTHORS                   clade; Panicoideae; Andropogoneae; Zea.					
TITLE                      1 (bases 1 to 604)					
JOURNAL                    Walbot V.					
COMMENT                   Maize ESTs from various cDNA libraries sequenced at Stanford University					
Contact: Walbot V					
Department of Biological Sciences					
Stanford University					
855 California Ave, Palo Alto, CA 94304, USA					
Tel: 650 723 2227 .					
Fax: 650 725 8221					
Email: walbot@stanford.edu					
Plate: 1117015 row: F column: 07.					
Location/Qualifiers					
1. .604					
/organism="Zea mays"					
/mol_type="mRNA"					
/db_xref="dbEST:946138f12.y1"					
/db_xref="taxon:4577"					
/clone_lib="1117 - Unigene V from Maize Genome Project"					
/notes="This library represents the unique genes found in					
for the fifth round of EST sequencing at Stanford University					
for the maize genome project. Sequences are present from					
library 946. Contigs were assembled using ZmDBAssembler					
and 2 representatives from each contig were selected for					
the Unigene set. All singlets were also selected."					
 ORIGIN					
Alignment Scores:					
Pred. No.:                   8 45e-66                   Length:                   604					
Score:                      997.50                   Matches:                   197					
Percent Similarity:        97.52%                   Conservative:              0					
Best Local Similarity:     97.52%                   Mismatches:              2					
Query Match:               34.28%                   Indels:                   3					
DB:                          14                           Gaps:                    2					
 US-10-030-884-14 (1-573) x CA831110 (1-604)					
Q	y		274	AlaaspPheAsnIleValGIYAlaAlaAlalysGlyGlyAlaAalaglyasp	293
D	b		6	GCCGACTTCTTCACATCGTGGTGGCGCCCAAGGAGCGGAGGACCGCGGGGAC	65
Q	y		294	GlucLylusGYAlaCySgLYGLYGLYGLYGLYHisSerProGlnAlaVal	313
D	b		66	GAGAGAGAGGGCGCATGC-----GGCGCGGAGGACACTCGCGCAGCGCGCGTC	119

library 946. Contigs were assembled using ZmDBAssembler and 2 representatives from each contig were selected for the Unigene set. All singlets were also selected."

ORIGIN

Alignment Scores:  
Pred. No.: 1.14e-64 Length: 630  
Score: 983.00 Matches: 202  
Percent Similarity: 96.19% Conservative: 0  
Best Local Similarity: 96.19% Mismatches: 5  
Query Match: 33.78% Indels: 3  
DB: 14 Gaps: 2

US-10-030-884-14 (1-573) x CA830783 (1-630)

QY 323 HisMetLeuValTrpSerSerAlaSerProValSerGlu-ArGAlaAlaValHisVa 342  
Db 7 CACATGCTCGTGGAGCTCCCTTTTCGCCCGCTCCGAGCGCG---CGCGTGCACGT 63  
QY 342 lPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTyrHis 362  
Db 64 CTTTCGCCGCGCGCGCGCTGACATGCCAGCTCTCCCAAGAGAGCCGAGCCCTACCA 123  
QY 362 pGluTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLys 382  
Db 124 CAGATACGGCGCGCGAGCTACTACAGCAGCAGACGAGAACGCGGAGCGCGCGCGACAA 183  
QY 382 sGlyGlyProThrLeuSerLysLeuGlySerLysSerThrAlaGlnLeuTyrProLysHis 402  
Db 184 GGGCGCGCGCGAGCTCTCGAAGCTGGGGTCCAACTCGACGGCGCAGCTGTATCCCAAGGA 243  
QY 402 pAspGlyGluGlyArgAlaAlaValAlaMetProProAlaSerValMetThrArgLe 422  
Db 244 CGACGGCGAGGGAGGAGCGAGCGCGG---CGATGCGCGCGCGCGGTGATCGCGGCT 300  
QY 422 uilleLeuileMetValTrpArgLysLeuileArgAsnProAsnThrTyrSerSerLeuile 442  
Db 301 CATCCTCATCATGCTGTGGAGGAAGCTGATCCGGAACCCCAACACTTACTCCAGCCTCAT 360  
QY 442 eGlyValValTrpSerLeuValSerTyrArgTrpGlyLysLeuMetProAlaAlaLeileal 462  
Db 361 CGCGCGTGTGTGCTCCTCGTATCCTACAGTGGGCACTCGAGATCCAGGATCATCGC 420  
QY 462 aArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLe 482  
Db 421 CGGTCGATTTGATCCTGTTCGAGCGGGTCTCGGGATGGCCATGTTCCAGCCTAGGCT 480  
QY 482 uPheMetAlaLeuGlnProArgIleleAlaCysGlyAsnLysLeuAlaAlaLeilealMe 502  
Db 481 GTTCATGGCGCTGCAGCGGAGGATCATCGCTGGGGAACAAGCTGGCGGCAATCCGAT 540  
QY 502 tGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLe 522  
Db 541 GGGGTCGCGTTCGTGCGAGGCGCGCGGTTCATGGCGCGGCTCCATCGCGTCCGTCT 600  
QY 522 uArgGlyValLeuLeuHisIleAlaile 531  
Db 601 GCGCGCGCTCCTCCTCCACATGCCATC 628

RESULT 13  
CG358058 726 bp DNA linear GSS 26-AUG-2003  
LOCUS CG358058  
DEFINITION CG358058  
CG358058 genomic survey sequence.  
ACCESSION CG358058  
VERSION CG358058.1 GI:34275325  
KEYWORDS GSS.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 726)

314 AlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerSerAlaSerPro 333  
Db 120 GCCGTGCGCGCGCGAGGAGGAGCTGCACATGCTCGTCTGAGCTTCAGCGCTCGCCC 179  
QY 334 ValSerGluArgAlaAlaValHisValPheGlyAlaGlyAlaAspHisAlaAspVal 353  
Db 180 GTCTCCGAGCGCGCGCGCTGCAGCTCTTCGCGCGCGCGCGCTGACCATGCCGAGCTC 239  
QY 354 LeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgThr 373  
Db 240 CTCGCCAAAGAGGCCGAGCTTACAGCAGTACGCGCGCGAGCTACAGCAGCAGGAGC 299  
QY 374 LysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsn 393  
Db 300 AAGACGGGAGCGCGCGCGAGCAAGCGCGCGCGCTGTCGAGCTGGGTCCCAAC 359  
QY 394 SerThrAlaGlnLeuTyrProLysAspAspGlyGlyArgAlaAlaAlaValAlaMet 413  
Db 360 TCGACGGCGAGCTACTATCCCAAGAGCAGCGCGAGGAGCGCGCGCG---GCCATG 416  
QY 414 ProProAlaSerValMetThrArgLeuileLeuileMetValTrpArgLysLeuileArg 433  
Db 417 CGCGCGCGGAGCTGATGACGGGCTCATCTCATCATGCTGTGGAGGAGCTGATCCGG 476  
QY 434 AnProAsnThrTyrSerSerLeuileGlyValValTrpSerLeuValSerTyrArgTrp 453  
Db 477 AACCCCAACACTTACTCAGCGCTCATCGCGCGTCTGCTGCTGCTGCTGCTGCTGCTG 536  
QY 454 GlyLysGluMetProAlaAlaIleleAlaArgSerIleSerIleLeuSerAspAlaGlyLeu 473  
Db 537 GGCATCGAGATCCAGCGATCATCCCGCGTGCATTCGATTCCTGTCGAGCGCGGCTC 596  
QY 474 GlyMet 475  
Db 597 GGGATG 602

RESULT 12  
CA830783 630 bp mRNA linear EST 12-DEC-2002  
LOCUS 1117011D11.y1 1117 - Unigene V from Maize Genome Project Zea mays  
DEFINITION cDNA, mRNA sequence.  
ACCESSION CA830783  
VERSION CA830783.1 GI:26558548  
KEYWORDS EST.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 630)  
AUTHORS Walbot V.  
TITLE Maize ESTs from various cDNA libraries sequenced at Stanford University  
JOURNAL Unpublished (1999)  
COMMENT Contact: Walbot V  
Department of Biological Sciences  
Stanford University  
855 California Ave, Palo Alto, CA 94304, USA  
Tel: 650 723 2227  
Fax: 650 725 8221  
Email: walbot@stanford.edu  
Plate: 1117011 row: D column: 11.  
Location/Qualifiers  
1. .630  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="taxon:4577"  
/clone\_lib="1117 - Unigene V from Maize Genome Project"  
/notes="This library represents the unique genes found in the fifth round of EST sequencing at Stanford University for the maize genome project. Sequences are present from

AUTHORS Whitelaw C.A., Quackenbush J., Van Aken, S., Utterback, T., Renick, A., Fraser C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T., Citek, R.W., Nummer, A., Robbins, D. and Lakey, N.

TITLE Consortium for Maize Genomics

JOURNAL Unpublished (2002)

COMMENT Other GSSs: OGI16K16TV  
Contact: Cathy Whitelaw

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: whitelaw@tigr.org  
Seq primer: TR  
Class: sheared ends.

FEATURES source  
1. 726  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clones="ZMMBMA0724C07"  
/clone\_lib="ZM 0.7 1.5 KB"  
/note="Vector: pBSK-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"

ORIGIN  
Alignment Scores:  
Pred. No.: 5,54e-63 Length: 726  
Score: 962.00 Matches: 192  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 33.06% Indels: 0  
DB: 29 Gaps: 0

US-10-030-884-14 (1-573) x CG358058 (1-726)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
Db 150 ATGATCACCGCGCTGGACCTTACACGTCTGACGGGGTGGCGCTGTACGTGCC 209  
Qy 21 MetThrLeuAlaTyrGlySerValArgTrrPArgIlePheThrProAspGlnCysSer 40  
Db 210 ATGACGCTGGCTAGCGCTCGCTGGTGGCGCATCTTACCGCGGACCAAGTCTCC 269  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 270 GGGATCAACCGCTTCTGGCGCTCTTGGCGTGGCGCTCTCTCTCTCCACTTCATCTCC 329  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
Db 330 ACCAACGACCCCTTCGCCATGAACTGGCTTCTTGGCGCGGACACGCTGCAGAGTG 389  
Qy 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
Db 390 GCCGTCTTCGGCGTCTGGCGCTGGCGCTTCCTCCCGCGCGCTCTCTCCCGCGCTCGGG 449  
Qy 101 LeuAspTrrPheIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
Db 450 CTCGACTGGAGCATCACGCTTCTTCCCTCTCCAGCTTCCCAACACGCTGCTCATGGCG 509  
Qy 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
Db 510 ATCCCGCTGCTGCGAGGCATGTACGGCGCTGCTGGCGCGCTCTCTCCCGCGCTCGGG 569  
Qy 141 ValValLeuGlnCysIleIleTrrPheThrLeuMetLeuPheLeuGluTyrArgAla 160  
Db 570 GTCGTCTTCCAGTGATCATCTGGTACACGCTTCAGCTCTTCTCTTCGAGTACCGGCC 629  
Qy 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
Db 630 GGGCGCGCGCTCTCTCGACCAAGTTCCTCCGACGGCGCGCGCTCCATGCTCTCTCTC 689  
Qy 181 ArgValAspSerAspValValSerLeuAlaArgGly 192

Db 690 CGCGTCGACTCCGACGTCTGCTCGTCCGCCAGGGGG 725

RESULT 14  
AY110024  
LOCUS Zea mays CL464\_3 mRNA sequence. 1221 bp mRNA linear HTC 17-OCT-2002  
DEFINITION Zea mays CL464\_3 mRNA sequence.  
ACCESSION AY110024  
VERSION AY110024.1 GI:21214085  
KEYWORDS HTC.  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE 1 (bases 1 to 1221)  
AUTHORS Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V. Design of Maize Mapping Project/DuPont Consensus Sequences for the Overgo Probes  
JOURNAL Unpublished (2002)  
AUTHORS 2 (bases 1 to 1221)  
Coe, E.H.  
TITLE Direct Submission  
JOURNAL Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA  
COMMENT If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, maizemap.org; ZmDB, www.zmdb.iastate.edu; TIGR, www.tigr.org; or NCBI, www.ncbi.nlm.nih.gov. When the source of the maize cDNA sequences is either Virginia Walbot, Stanford or Pat Schnable, Iowa State, then clones may be requested from ZmDB: www.zmdb.iastate.edu.

FEATURES  
Location/Qualifiers  
1. 1221  
/organism="Zea mays"  
/mol\_type="mRNA"  
/db\_xref="MaizeDB:632107"  
/db\_xref="taxon:4577"  
/clone\_lib="Maize Mapping Project/DuPont Consensus Library"  
/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed DuPont contigs; this resource was assembled by DuPont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"

ORIGIN  
Alignment Scores:  
Pred. No.: 1.14e-60 Length: 1221  
Score: 936.00 Matches: 212  
Percent Similarity: 63.75% Conservative: 35  
Best Local Similarity: 54.50% Mismatches: 75  
Query Match: 32.16% Indels: 66  
DB: 11 Gaps: 12

US-10-030-884-14 (1-573) x AY110024 (1-1221)

Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
Db 142 ATGATCACCGCGACGACTTACACGCTGATCAGCGCATGTCCTGCTGTACGTGCC 201  
Qy 21 MetThrLeuAlaTyrGlySerValArgTrrPArgIlePheThrProAspGlnCysSer 40  
Db 202 ATGATCTCGCTAGCATCCGTCAGTGGTGGCGCATCTTACCGCGGACCAAGTCTCG 261  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 262 GGGATCAACCGCTTCTGGCGCTCTTCGCGCGCTCTCTCTCTCCACTTCACATCANN 321  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
Db 611 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80

322 NNNNNNNNNNNNNNNNNNNAAACATCCGGTTCATCCGCGCGCAACCCGCGCCTCCGGCTG 381

81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96

382 ATGTCCTCGCGCTGCTCACTGATGAGACTACTCTCCGCGCGGGCTGC-----432

97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerIleuSerThrLeuProAsnThr 116

433 -----CTCAGATGGACCATCAGCTCTTCTCCCTGTCGACGCTGCCAACACG 480

117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136

481 CTGTCGATGGGCATCCGCTGCTCAAGGCATGTACGGCACTTCC---GGCAGCCTC 537

137 MetValGlnValValValLeuGlnCysIleIleTyrTyrThrLeuMetLeuPheLeuPhe 156

538 ATGTCGCAGATCGTGTGCTCCAGTCATCATCTGTCACACGCTGATGCTGTTCATGTTTC 597

157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAsp-----GlyAlaAla 174

598 GAGTACCGCGCGCCAGGATCTCATCCAGCAGGTTCCTCCGACACGCGCGGCGGCATC 657

175 AlaSerIleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAsp 193

658 GCCTCCATCGTG-----GTGGACCCCGACGTGTGTCTGTCGACGCGGCACACGAC 708

194 ---ValGluLeuGluAlaGluProAspGlyValAlaGlyAlaValSerSerArg 212

709 GCCATCGACGCGAGGCCGAGGTGAAG-----735

213 GlyGlyAspAlaGlyArgValArgValThrValArgIleSerThrSerSerArgSerGlu 232

736 -----GAGACGGCAAGATACACGTCCGTCGCGCTCCAACCGTCGCTCGGAC 789

233 AlaAlaCysSerHisSerHisSer---GlnThrMetGlnProArgValSerAsnLeuSer 251

790 ATCTACTCTCCGCGGTCCTAGGGGTTCTCCAGCACACCGCGCGCCAGCACTTGACC 849

252 GlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPhe 271

850 AACCGCGAGATCTACTCGCTGTCAGTCGTCAGAGAACCCCAACGCGCGGGGCTCCAGCTTC 909

272 AsnHisAlaAspPhePheAsnIleValGlyAlaAlaAlaGlyGlyGlyAlaAla 291

910 AACCAACCCACTTCTACTCTCATGTCGCGCGCAGTCTCCAACTTCGCGCGCGGGACGG 969

292 -----GlyAspGluIleGlyGly 297

970 TTCGGCTCGCACGCGGCGCCACCGCCAGSCGTCCAACTACGAGGAGGACCGCAGGCG 1029

298 AlaCysGlyGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaValAlaValProAla 317

1030 AAGCGCAACAAGTACGCGCCAGTACCGCGCGCCCAACCCGCGCATGCGCGGACGCGCGC 1089

318 Lys-----ArgIleAspLeuHis 323

1090 AAGGGCTCAAGAGGGCGGCATGGCAGGCCAAGGGCGAGGACCGCAGGACCTTACAC 1149

324 MetLeuValTrpSerSerAlaSerProValSerGluArgAlaAlaValHisValPhe 343

1150 ATGTCGTGTGAGCTCCAGCGGTGCGCCGTGTCCGAC-----GTGTTCC 1194

344 GlyAlaGlyAlaAspHisAlaAsp 352

1195 GGCAATGGCGCGCCGAGTACACGAC 1221

RESULT 15	CC371166	972 bp	DNA	linear	GSS 19-MAY-2003
LOCUS	CC371166				
DEFINITION	PHERSK16TB ZM 0.6_1.0 KB Zea mays genomic clone ZMMBPa523D07, genomic survey sequence.				
ACCESSION	CC371166				
VERSION	CC371166.1				GI:30844783

```
Db 484 CCAGCGATCATCGCCCGTGCATTTTCATCTGTCGACGCGGTCCTCGGGATGCGCATG 543
Qy PheSerLeu----- 480
Db 544 TTCAGCCTAGGTGTGTGTGTCTTACTTGTACGGCGTCCCGTGTCTCTGCTGACCCGCTG 603
Qy 481 -----GlyLeuPheMetAlaLeu 486
Db 604 CAATCGAGCTCTGAACCGCGCCCTTGCCTCTGTTCAGGCTGTTCATGGGCTG 663
Qy 487 GlnProArgIleIleAlaCysGlyAsnIlyLeuAlaIleAlaMetGlyValArgPhe 506
Db 664 CAGCCGAGGATCATCGGCTGCGGAAACAAGCTGGCGCCATCGCGATGGCGTCCCGGTTT 723
Qy 507 ValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuArgGlyValLeu 526
Db 724 GTCGAGGCCCGCGGTTCATGCGCGCTCCATCGCGTCTGCGGCGTCTC 783
Qy 527 LeuHisIleAlaIleVal----- 532
Db 784 CTCACATCGCCATCGT-CCAGGTACGCGCCCTGTGCTTGAACGTAAACGCTGGAGTTGT 842
Qy 533 -----GlnAlaAlaLeuProGln 538
Db 843 GATGAACAGTAGGACCAACATTCGTTTCCCTTCTTCTCCAGGCTGCTCTGCTCAG 902
Qy 539 GlyIleValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThr 558
Db 903 GGGATCGTGGCGTTCGTGTTCCGCAAGGAGTACGGCGTTTCATCCCGACATCCTGAGCACA 962
Qy 559 AlaTyrGly 561
Db 963 GCGTATGGT 971
```

Search completed: March 3, 2004, 12:16:03  
Job time : 4006 secs



222 CGCGATGTGGATCACCTA-----TGC CGCGATCTACTCTCGCCACACAGCGTGTTCAA 275

340 ValHisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGln 359

276 AGCCACGCCGCGAGGATGCCGGGTCTCACCTCACCTCGCTCGCCAACTATCGCG 335

360 AlaTyAspGluTyGlyArgAspAspTySerSerArgThrLysAsnGlySerGlyGly 379

336 GCTAGGTCGCCGATCCTCGGCAGCGTCTGGGTGAGGCGCGCTCAACCTCGCTCGCT 395

380 AlaAspLysGlyGlyProThrLeuSerLysLeuGlySerLysSerThrAlaGlnLeuTy 399

396 -----GGCGGTCTCTATCGCTCGCGGTCTGGTACTGATGAC-----CCCGTTCG 440

400 ProLysAspAspGly-GluGlyArgAlaAlaAlaValAlaMetProProAlaSerValMe 419

441 CCTGCTGATCTGGAGCGGTGAAAAAGCCGCCGCCGGTGAACAGCGGTTCTACGCT 500

419 tThrArgLeuLeuLeuLeuMetValTrpArgLysLeuLeuLeuLeuLeuLeuLeuLeu 439

501 GGCAATGCTCGCGGTGGTGTGATG---TGCGGTGC---GTGMAAACCAATCGCTCGGG 554

439 rSerLeuLeuGlyValValTrpSerLeuValSerTy-ArgTrpGlyLeuGluMetProAl 459

555 CCGCTGCTTGGGGTGGTGTCTTCGCGCATC-----GGCAATTAATATGCCGGA 602

459 aIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPhe 479

603 CTGCTGCTGGGTGATCGATCAACCGCTGGCGCTGGCCGCCCGCGCGGTGTTCCT 662

479 rLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAla 499

663 CACCGGGGTGATCTCGTGGCGCGAAACTCGACGTC-----AATCGCTGATCGC 713

499 aIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerIleAl 519

714 CACCTCCACCATCGTGAAACTCTGCTGACAGCGGTTTATTCCCTGGGTCTGGTGATG 773

519 aValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaLeuProGlnG 539

774 ACTTGGGCTGCACGGTTCATTACCGCATCTGATGATTCCTGCTCGCGCAGG 833

539 yIleValProPheValPheAlaLysGluTyGlyValHis---ProAsp 554

834 CTCTTTTGGCGTGTCTTCGGCAACCGCTTCGGGTCCAGTCTCCGAT 882

## RESULT 2

```

US-09-335-409-1
; Sequence 1, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-335-409-1

```

Alignment Scores:	
Pred. No.:	0.047
Score:	152.50
Length:	68750
Matches:	106
Conservative:	43
Percent Similarity:	29.68%

Best Local Similarity:	21.12%	Mismatches:	164
Query Match:	S.24%	Indels:	190
DB:	3	Gaps:	14

US-10-030-884-14 (1-573) x US-09-335-409-1 (1-68750)

Qy	118	ValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMet	137
		:::	:::
Db	28213	GTCTGTCGGGGGTCTTCCTCGAAGATGGGTGAGGCTCGTGGCGGCGCG	28266
Qy	138	ValGlnValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuGlu	157
Db	28266	-----	28266
Qy	158	TyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIle	177
Db	28267	---CGGCGCGCTGATCGAGGGGCTCTCGCGCGCGCGCGCGATGGTGTCTCGTCGGAGC---	28320
Qy	178	ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu	197
Db	28320	-----	28320
Qy	198	AlaGluProAspGlyValAlaGlyAlaGlyAlaVal	209
Db	28321	GCCGAGCGCGAGGTGCGCGCGCGCGGTGCGCGCACGCGCGTCTCGTTCGATCGCGCG	28380
Qy	210	-----SerSerArgGlyAspAlaGlyArg	220
	:::		
Db	28381	GTCTAATGGCGCGGACAGTGTGTATCGGGCGCTGGAGCAAGCGGTGAGCGGATCGC	28440
Qy	221	ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys-SerHis	239
	:::	:::	:::
Db	28441	GGCGGGTTCGGCGCGCGCGCGCGCACCAAGCGGTGCTGCTCGCACGGCTCCCA	28500
Qy	239	sSerGlnThrMetGlnPro-----ArgValSerAsnLeuSerGlyVa	253
Db	28501	CTCGCGCTGATGAACCATCTGTGAGGAGTTCGGCGGTGGCGGC-GTCTGTCACGT	28559
Qy	253	lGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHi	273
Db	28560	ACCGCGCGCAACGCTTCGTGCTGTCGAGCAACCTGAGGGGAAGTGTGTCG- --GGACG	28616
Qy	273	sAlaAspPheAenIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAs	293
Db	28617	AGCTGAGCGCGCGGGTACTGGGTGGCGCACGTGCGGAGCGGTGCGCTTCGCGGACG	28676
Qy	293	pGluGluLysGlyAlaCysGlyGlyGlyGlyHisSerProGlnProGlnAlaVa	313
Db	28677	GGGTGAAGCGCTGCACGAAGCGGTGCGGCGCACGTTCTCGAAGTGGGCCGGAAGCCGA	28736
Qy	313	lAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPr	333
Db	28737	CGCTGTCT-----	28743
Qy	333	oValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAspVa	353
Db	28744	-----CGGGCTGTTCCAGCCTGCTCCGAGCGCGGAGCGGACGTCTGGCTGGCGT	28793
Qy	353	lLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgTh	373
Db	28794	CGTTGCGCGC-----CGGCG	28808
Qy	373	rLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAs	393
Db	28809	GCAGAGAGCTCGGGGGTGTCTGAGCGGCTGGCAGGCTGTGGCGCGCGCGCGGCTCGG	28868
Qy	393	nSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArgAlaAlaAlaValaMe	413
Db	28869	TCAGCTGGCGGCGCTTCTCCACGCGCTGGCGGGTGCCTGCGCTCGCGACTATCGGT	28928
Qy	413	tProProAlaSerValMetThr-----	420
Db	28929	GGCAGGCGGACGCGGTACTGCGCGCATCGACCTGACAGCGCTGCCACGACGCGCGG	28988

QY 421 -----ArgLeuIleLeuLeu-MetValTTPArgLysLeuLeuLeuArgAsn----- 434  
 Db 28989 ATCCGACCAAGGCTGGTCTATCCCGTGGAGTGGCGGAGATACCTCGAGCGCTCCAGA 29048  
 QY 434 ----- 434  
 Db 29049 AATCAGAGGAGCGAGCGCGGAGCTGGCTGGTATTGGCGGATAAGGTGGAGTCGGCG 29108  
 QY 435 -----ProA 436  
 Db 29109 AGCGGTGCTGCACGCGCTGCGACACGTGGACTTCCATGCGTGTGCTCCATCGCGCGG 29168  
 QY 436 snThrTySerSerLeuIleGlyValVal-----TTPs 447  
 Db 29169 CAGAGACATCCGCGACCGCCGAGCTGGTGACCGAGGTGCGCGCGGTGCGAAGCGATTGGC 29228  
 QY 447 erLeuValSerTyArgTTPGlyIleGluMetProAlaIleIleAlaAArgSerIleSerI 467  
 Db 29229 AGGTAGTCTCTACTGTGGGTCTGGACGCGCTGCGGTGCGGAGCGCTCGATCGATG 29288  
 QY 467 leLeuSerAspAlaGly-----LeuGlyMetAlaMetPheSerL 480  
 Db 29289 AGATCGCGGACGCGACCGCTGCTGTACCGCGCGGTGCTCGGTGGCTGCTGCTGCTG 29347  
 QY 480 euGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIaI 500  
 Db 29348 -----AGCACCGTGTCTGTTCGCCCGCGACTCTGGGTGCG 29381  
 QY 500 leAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaV 520  
 Db 29382 TGACCGGGGGGCGATCGATCGTTGGCGACGAGCCTGGATCGCCCTTGTTCAGCGCGCT 29441  
 QY 520 al 520  
 Db 29442 TA 29443

## RESULT 3

US-09-568-102-1  
 ; Sequence 1, Application US/09568102  
 ; Patent No. 6346404  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schupp, Thomas  
 ; APPLICANT: Ligon, James  
 ; APPLICANT: Molnar, Istvan  
 ; APPLICANT: Zirkle, Ross  
 ; APPLICANT: Cyr, Devon  
 ; APPLICANT: Goerlach, Joern  
 ; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
 ; FILE REFERENCE: 4-30582A  
 ; CURRENT APPLICATION NUMBER: US/09/568,102  
 ; CURRENT FILING DATE: 2000-05-10  
 ; PRIOR APPLICATION NUMBER: 09/335,409  
 ; PRIOR FILING DATE: 1999-06-17  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 68750  
 ; TYPE: DNA  
 ; ORGANISM: Sorangium cellulosum  
 US-09-568-102-1

## Alignment Scores:

Pred. No.: 0.047 Length: 68750  
 Score: 152.50 Matches: 106  
 Percent Similarity: 29.68% Conservative: 43  
 Best Local Similarity: 21.12% Mismatches: 164  
 Query Match: 5.24% Indels: 190  
 DB: 4 Gaps: 14

US-10-030-884-14 (1-573) x US-09-568-102-1 (1-68750)

QY 118 ValMetGlyIleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThrLeuMet 137

Db 28213 GTGCTGGCGGGGGTCTTCCTCGTGAAGATGGGTGAGGCTCGTGGCGCGCG----- 28266  
 QY 138 valGlnValValValLeuGlnCysIleIleTTPTyThrLeuMetLeuPheLeuPheGlu 157  
 Db 28266 ----- 28266  
 QY 158 TyArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIle 177  
 Db 28267 ---CGGCGGCTGATGACGGGGCTCTCGCGCGCGCGCGATGGTGTCTCGCTCGGAGC--- 28320  
 QY 178 ValSerPheArgValAspSerAspValValSerLeuAlaAArgGlyAspValGluLeuGlu 197  
 Db 28320 ----- 28320  
 QY 198 AlaGluProAspGlyValAlaGlyAlaGlyAlaVal----- 209  
 Db 28321 GCGGAGGCGGAGGTGGCGCGCGCGGTGCGCGCGCGCGCGCGGTGCTGCTGATCGCGCG 28380  
 QY 210 -----SerSerArgGlyGlyAspAlaGlyArg-----ValArg 220  
 Db 28381 GGTCAATGGCGCGAGCAGGTGCTGATCGCGCGCTGGAGCAAGCGGTGCGAGCGATCGC 28440  
 QY 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys-SerHis---SerHi 239  
 Db 28441 GCGGGGTTCG 28500  
 QY 239 sSerGlnThrMetGlnPro-----ArgValSerAsnLeuSerGlyVa 253  
 Db 28501 CTCGCCGCTGATGGAACCGATCTCGAGGAGTTCGGGGGTGGCGGC-GTCCGTGACGT 28559  
 QY 253 lGluIleTySerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHi 273  
 Db 28560 ACCGCGCGCAAGCGTTTCGTGTGAGCAACTGAGCGGGAAGTGTGTCG---GGACG 28616  
 QY 273 sAlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAs 293  
 Db 28617 AGCTGAGCGCGCGGGGTACTGGTGGCGCACGTGCGGAGCGGTGCGCTTCGCGGACG 28676  
 QY 293 pGluGlnLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaVa 313  
 Db 28677 GGGTGAAGCGCTGCACGAGCGCGTGGCGGACGCTTCGTAAGTGGCGCGCGAGCGCGA 28736  
 QY 313 lAlaValProAlaLysArgLysAspLeuHisMetLeuValTTPSerSerSerAlaSerPr 333  
 Db 28737 CGCTGCT----- 28743  
 QY 333 oValSerGluArgAlaAlaValHisValPheGlyAlaGlyAlaAspHisAlaAspVa 353  
 Db 28744 -----CGGGCTGTTGCCAGCTGCTCGCGAGCGGAGCGGCGGTGCTGCTGGCGT 28793  
 QY 353 lLeuAlaLysGlyAlaGlnAlaTyAspGluTyArgAspAspTySerSerArgTh 373  
 Db 28794 CGTTCGCGC-----CGGCG 28808  
 QY 373 rLysAsnGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAs 393  
 Db 28809 GCGAGGAGCTCGCGGGGTGCTCGAGGCGCTGGCGGCTGGCGCGCGCGCGCGCTCGG 28868  
 QY 393 nSerThrAlaGlnLeuTyProLysAspAspGlyGluGlyAlaAlaAlaValAlaMe 413  
 Db 28869 TCAGTGGCGCGGCTTTCCTCCACCGGTGGCGCGGCTGCGCTGCGACCTATCCGT 28928  
 QY 413 tProProAlaSerValMetThr----- 420  
 Db 28929 GGCAGCGCGCAGCGGTACTGGCCCGACATCGAGCGCTGACAGCGCTGCCACGCGCGCGG 28988  
 QY 421 -----ArgLeuIleLeuIle-MetValTTPArgLysLeuIleArgAsn----- 434  
 Db 28989 ATCCGACCAAGGCTGGTCTTATCGGTGGATGTCGCGCGGAGATACCTCGCAGCCTCCAGA 29048  
 QY 434 ----- 434

Db 29049 AATCAGAGGAGCGAGCGCGGAGCTGGCTATTGGCGGATAAGGGTGGAGTCGGCG 29108  
Qy 435 -----ProA 436  
Db 29109 AGCGGTCTCGTACGCGTGTTCACACGTGGACTTCCATGCTGCTCCATGCGCGG 29168  
Qy 436 snThrTyrSerSerLeuIleGlyValVal-----Trps 447  
Db 29169 CAGAGACATCCGCGAGCGCGGAGCTGGTGACCGAGGCTGCGCGGCGTGAAGCGATTGC 29228  
Qy 447 erLeuValSerTyrArgTTPGlyVileGluMetProAlaIleIleAlaIleAlaIleAlaI 467  
Db 29229 AGGTAGTGTCTACCTGTGGGTCTGGACCGCGTCTGGTGGTGGTGGTGGTGGTGGTGG 29288  
Qy 467 leLeuSerAspAlaGly-----LeuGlyMetAlaMetPheSerL 480  
Db 29289 AGATCGCGCAGCGGACCGCTGCTGTACCGCGCGGCTGGCTGGCTGGCTGGCTGGCTGGCT 29347  
Qy 480 euGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaI 500  
Db 29348 -----AGCACCGTGTCTTGTTCGCCCGGACTCTGGGTGCG 29381  
Qy 500 leAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaIleAlaIleAlaI 520  
Db 29382 TGACCGCGGGGCGCATGCTGTGGCGGAGCGCTGGCGATCGCCCTTGTTCAGCGCGGCGT 29441  
Qy 520 al 520  
Db 29442 TA 29443  
RESULT 4  
US-09-567-969-1  
; Sequence 1, Application US/09567969  
; Patent No. 6355457  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Geerlach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/567,969  
; PRIOR FILING DATE: 2000-05-10  
; PRIOR FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 68750  
; TYPE: DNA  
; ORGANISM: Sorangium cellulosum  
US-09-567-969-1  
Alignment Scores:  
Pred. No.: 0.047 Length: 68750  
Score: 152.50 Matches: 106  
Percent Similarity: 29.68% Conservative: 43  
Best Local Similarity: 21.12% Mismatches: 164  
Query Match: 5.24% Indels: 190  
DB: 4 Gaps: 14  
US-10-030-884-14 (1-573) x US-09-567-969-1 (1-68750)

Qy 158 TyArgAlaAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIle 177  
Db 28267 ---CGGGCGGCTGATGACGGGCTCTCGCGGCGCGGCGATGGTCTCGCTCGGAGC--- 28320  
Qy 178 ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu 197  
Db 28320 ----- 28320  
Qy 198 AlaGluProAspGlyValAlaGlyAlaGlyAlaVal----- 209  
Db 28321 GCCGAGCGCGAGGTGGCGCGCGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 28380  
Qy 210 -----SerSerArgGlyGlyAspAlaGlyArg-----ValArg 220  
Db 28381 GGTCAATGGCGCGAGCAGGTGGTGTATCGCGGCGGTGGAGCAAGCGGTGGCAGCGCATCG 28440  
Qy 221 ValThrValArgGlySerThrSerSerArgSerGluAlaAlaCys-SerHis---SerHi 239  
Db 28441 GCGGGGTTTCG 28500  
Qy 239 sSerGlnThrMetGlnPro-----ArgValSerAsnLeuSerGlyVa 253  
Db 28501 CTCGCCGCTGATGGAAACCGATGCTGGAGGAGTTCGGCGCGGTGGCGGC-GTCGGTGACGT 28559  
Qy 253 lGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHi 273  
Db 28560 ACCGCGCGCCAGCGCTTTCGCTGGTGAGCAACCTGAGCGGGAAGGTGGTCGC---GGACG 28616  
Qy 273 sAlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAs 293  
Db 28617 AGCTGAGCGCGCGCGGCTACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 28676  
Qy 293 pGluGluLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 313  
Db 28677 GGGTGAAGCGCTGCACGAAGCGCGGTGGCGGCGACGTTCTCGAAGTGGGCGCGCGCGCG 28736  
Qy 313 lAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPr 333  
Db 28737 CGTGTCT----- 28743  
Qy 333 oValSerGluArgAlaAlaValHisValPheGlyAlaGlyAlaAspHisAlaAspVa 353  
Db 28744 -----CGGGCTGTTGCCAGCCTGCTGCGGAGCGGAGCGCGCGCTGCTGGCGT 28793  
Qy 353 lLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyr-SerSerArgTh 373  
Db 28794 CGTTCGCGC-----CGGCG 28808  
Qy 373 rLysAsnGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAs 393  
Db 28809 GCGAGGAGGTGCGGGGTGCTCGAGCGCTGGGCGAGGCTGTGGCGCGCGCGCGCGCTCG 28868  
Qy 393 nSerThrAlaGlnLeuTyrProLysAspAspGlyGlyArgAlaAlaValAlaMe 413  
Db 28869 TCAGTGGCGCGGCTTCTCCCGCGCGGCGGCTGGCGCTGGCGCTATATCCGT 28928  
Qy 413 tProProAlaSerValMetThr----- 420  
Db 28929 GGCAGCGCGCAGCGGTACTGGCGCGCGACATCGAGCCTGACAGCGCTGCCACGCGCGCG 28988  
Qy 421 -----ArgLeuIleLeuIle-MetValTrpArgLysLeuIleArgAsn----- 434  
Db 28989 ATCCGACCCAAAGGTGGTCTATCGCGTGAGCTGGCGGAGATACCTCGCAGCGCTCAGA 29048  
Qy 434 ----- 434  
Db 29049 AATCAGAGGAGCGAGCGCGGAGCTGGTGGTATTGGCGGATAAGGTGGAGTGGCGCG 29108  
Qy 435 -----ProA 436  
Db 29109 AGCGGTCTCGTACGCGCTGTTCGACACGTGGACTTCCATGCTGCTCCATGCGCGCGG 29168  
Qy 436 snThrTyrSerSerLeuIleGlyValVal-----Trps 447

Db 29169 CAGAGACATCCGCGACCGCGAGCTGGTACCGAGGCTCGCGGCTCGAAGCGATTGGC 29228  
QY 447 erLeuValSerTyrArgTTrpGlyLeuMetProAlaIleAlaAraSerIleSerI 467  
Db 29229 AGGTAGTCTCTACTGTGGGCTTGACCGCCGTCGTCGGTGGCGGCGTCGATCATG 29288  
QY 467 leLeuSerAspAlaGly-----LeuGlyMetAlaMetPheSerL 480  
Db 29289 AGATCGCGGACGACCGCTGCTACCGCGCGGCTCGCTCGGCTTGGCTGCTTCTG 29347  
QY 480 euGlyLeuPheMetAlaLeuGlnProArgIleAlaCysGlyAsnLysLeuAlaI 500  
Db 29348 -----AGCACCGGTGCTTGTTCGCCCGACTCTGGTTCG 29381  
QY 500 leAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaV 520  
Db 29382 TGACCCGGGGGATGATCGTTGGCGACGAGCTCGGATCGCCCTTGTGACGGCGGT 29441  
QY 520 al 520  
Db 29442 TA 29443  
RESULT 5  
US-09-568-480-1  
; Sequence 1, Application US/09568480  
; Patent No. 6355458  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Goslach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/568,480  
; PRIOR FILING DATE: 2000-05-10  
; PRIOR APPLICATION NUMBER: 09/335,409  
; PRIOR FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 68750  
; TYPE: DNA  
; ORGANISM: Sorangium cellulosum  
US-09-568-480-1  
Alignment Scores:  
Pred. No.: 0.047 Length: 68750  
Score: 152.50 Matches: 106  
Percent Similarity: 29.68% Conservative: 43  
Best Local Similarity: 21.12% Mismatches: 164  
Query Match: 5.24% Indels: 190  
DB: 4 Gaps: 14  
US-10-030-884-14 (1-573) x US-09-568-480-1 (1-68750)  
QY 118 ValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMet 137  
Db 28213 GTGGTGGCGGGGTGTTCTCGCTGGAGATGGGTGAGGCTCGTGGCGCGCG----- 28266  
QY 138 ValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuGlu 157  
Db 28266 ----- 28266  
QY 158 TyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaSerIle 177  
Db 28267 ---CGGGCGGCTGATGACAGGGGCTCTCGGGGGCGGCGGATGTTGCTCGGAC--- 28320  
QY 178 ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu 197  
Db 28320 ----- 28320

QY 198 AlaGluProAspGlyValAlaGlyAlaGlyAlaVal----- 209  
Db 28321 GCGGAGGCGGAGGTGGCGGCGCGGTGGCGCCACCGCGCGTGGTGTGATCGCGCG 28380  
QY 210 -----SerSerArgGlyGlyAspAlaGlyArg-----ValArg 220  
Db 28381 GGTCAATGGCGCGGACGAGTGTGATCGCGGCGTGGAGCAAGCGGTGCGAGCGCATCGC 28440  
QY 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys-SerHis---SerHi 239  
Db 28441 GCGGGGTTCGCGCGCGCGCGCGCGCCACCAAGCGGTGATGCTCTCGACCGCTCCCA 28500  
QY 239 sSerGlnThrMetGlnPro-----ArgValSerAsnLeuSerGlyVa 253  
Db 28501 CTCGCCGCTGATGAAACCGATGCTCGAGAGTTTCGGCGGGTGGCGGC-GTCGGTCACT 28559  
QY 253 lGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHi 273  
Db 28560 ACCGGCGGCAAGCGTTTCGTGTGTGACCAACCTGAGCGGAAGGTGTGCG---GGACG 28616  
QY 273 sAlaAspPhePheAsnIleValGlyAlaAlaLysGlyGlyGlyAlaAlaGlyAs 293  
Db 28617 AGCTGAGCGCGCGGGGTACTGGGTGGCGCACGTGCGGAGCGGTGCGCTTGGCGGACG 28676  
QY 293 pGluGlnLysGlyAlaCysGlyGlyGlyGlyHisSerProGlnProGlnAlaVa 313  
Db 28677 GGGTGAAGGCGCTGCACGAAGCGGTGCGGCGCCTTCGTGAAGTGGGCGCGAGCCGA 28736  
QY 313 lAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPr 333  
Db 28737 CGCTGCT----- 28743  
QY 333 oValSerGluArgAlaAlaValHisValPheGlyAlaGlyAlaAspHisAlaAspVa 353  
Db 28744 -----CGGGCTGTTCAGCCTCCCTCCGAGCGGAGCGGACCGCTGTGGCGT 28793  
QY 353 lLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgTh 373  
Db 28794 CGTTGCGGCG-----CGGCG 28808  
QY 373 rLysAsnGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAs 393  
Db 28809 GCGAGAGGCTGCGGGGTGCTCGAGGCGCTGGGCGCGCTGGGCGCGCGCGCTCGG 28868  
QY 393 nSerThrAlaGlnLeuTyrProLysAspAspGlyGlyArgAlaAlaValAlaMe 413  
Db 28869 TCAGCTGGCGGGCGCTTCCCGCGGCTGGCGGCGGTGCGCTGCGGACCTATCCGT 28928  
QY 413 tProProAlaSerValMetThr----- 420  
Db 28929 GGCAGCGGACGCGGTACTGGCGCGACATCGAGCCTGCGCCACGCGCGCGCG 28988  
QY 421 -----ArgLeuIleLeuIle-MetValTrpArgLysLeuIleArgAsn----- 434  
Db 28989 ATCCGACCAAGGCTGGTTCATCGGTGGACTGGCGGAGATACCTCGAGCTCCAGA 29048  
QY 434 ----- 434  
Db 29049 AATCAGAGAGCGAGCGCGGAGCTGGCTGTATTGGCGGATAAGGGTGGAGTCCGCG 29108  
QY 435 -----ProA 436  
Db 29109 AGCGGCTGCTGCGAGCGGTGTCGACACGTGACTTCCATGCGTGTGCTCATCGCGCG 29168  
QY 436 snThrTyrSerSerLeuIleGlyValVal-----Trps 447  
Db 29169 CAGAGACATCCGCGACCGCGGCTGTGACCGAGGCTGCGCGGTGCGAGCGATTGGC 29228  
QY 447 erLeuValSerTyrArgTTrpGlyLeuMetProAlaIleAlaArgSerIleSerI 467  
Db 29229 AGGTAGTCTCTACTGTGGGCTTGAGACGCGCGCTGCTCGGTGGCGGCGGTGCTGATG 29288

```

Qy   467 leUeSerAspAlaGly-----LeuGlyMetAlaMetPheSerL 480
      ::::| | | | | | | | | | | | | | | | | | | | | |
Db   29289 AGATCGGCAGCGACCCGGTCGTCTACGCCCGCGTGTGGCTTCGGTTCTCG- 29347
      : : : : : : : : : : : : : : : : : : : : : :
Qy   480 euGlyLeuPheMetAlaLeuGlnProArgIleAlaCysGlyAsnLysLeuAlaAlaI 500
      : : : : : : : : : : : : : : : : : : : : : :
Db   29348 -----AGCACCGTGCTCTGTTCCGCCCGACTCTGGGTCTG 29381
      : : : : : : : : : : : : : : : : : : : : : :
Qy   500 leaLaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerlleAlav 520
      : : : : : : : : : : : : : : : : : : : : : :
Db   29382 TGACC CGGGGGCATCATCGTTGGCAGCAGCCTCGCATCCCTTGTTCAGCGCGCGT 29441
      : : : : : : : : : : : : : : : : : : : : : :
Qy   520 al 520
      : :
Db   29442 TA 29443

RESULT 6
US-09-568-486-1
; Sequence 1, Application US/09568486
; Patent No. 6355459
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Molnar, James
; APPLICANT: Logonar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,486
; PRIORITY FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 68750
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-568-486-1

Alignment Scores:
Pred. No.:          0.047           Length:       68750
Score:             152.50         Matches:        106
Percent Similarity: 29.68%        Conservative:    43
Best Local Similarity: 21.12%     Mismatches:     164
Query Match:       5.24%          Indels:         190
DB:                4              Gaps:           14

US-10-030-884-14 (1-573) x US-09-568-486-1 (1-68750)

Qy   118 ValMetGlyIleProLeuLeuArgGlyMetTyrglyAlaSerSeraGlyThrLeuMet 137
      : : : : : : : : : : : : : : : : : : : : : :
Db   28213 GTGCGTGGCGGGGTGTCTTCGCTGGAAGATGGGTGAGGCTCGTGGCGCGCG- 28266
      : : : : : : : : : : : : : : : : : : : : : :
Qy   138 ValGlnValValLeuLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheGlu 157
      : : : : : : : : : : : : : : : : : : : : : :
Db   28266 ----- 28266

Qy   158 TyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaSerIle 177
      : : : : : : : : : : : : : : : : : : : : : :
Db   28267 ---CGGCGCGCTGATGACAGGGGCTCTCGGCGGGCGGCGATGTGCTCGGAGC--- 28320
      : : : : : : : : : : : : : : : : : : : : : :
Qy   178 ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu 197
      : : : : : : : : : : : : : : : : : : : : : :
Db   28320 ----- 28320

Qy   198 AlaGluProAspGlyValAlaGlyAlaGlyAlaVal----- 209
      : : : : : : : : : : : : : : : : : : : : : :
Db   28321 GCCGAGCGCGAGGTGGCGCGCGCGGTGGCGCGCGCGCGTCCGTCGATCGCGCG 28380
      : : : : : : : : : : : : : : : : : : : : : :
Qy   210 -----SerSerArgGlyGlyAspAlaGlyVarg-----ValArg 220
      : : : : : : : : : : : : : : : : : : : : : :

```

28381	Db	GETCAATGGCCGGACGAGGTGGTTCATCGCGCGGTGGAGCAAGCGGTGCAGGCGATCGC	28444
221	QY	ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys-SerHis----	239
28441	Db	GGCGGGGTTCGGCGCGCGCGCGCCACCAAGCGGTGCATGCTCGCACGGTCCCCA	28500
239	QY	sSerGlnThrMetGlnPro-----ArgValSerAsnLeuSerGlyVa	253
28501	Db	CTTCGGCGCTGATGGAACCGATCTGGAGGAGTTCGGCGGGTGGCGGC-GTCCGTGACGT	28559
253	QY	IgluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHi	273
28560	Db	ACCGCGCGCCAAAGCGTTTCGTGTGTGTGACCACTTCAGCGGGAAGTGGTGC-	28616
273	QY	salAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAs	293
28617	Db	AGCTGAGCGCGCGGGGTACTGGGTGCGGCACGTCTGGGAGCGGTTCGTCGCGACG	28676
293	QY	pGluGlnLysGlyAlaCysGlyGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaVa	313
28677	Db	GGGTGAAGCGCTGCACGAAGCGGTGGGCGACGTCTGCTCAAGTGGCGCCGAAGCCGA	28736
313	QY	lalaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPr	333
28737	Db	CGCTGCT-----	28743
333	QY	oValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspVa	353
28744	Db	-----CGGGCTGTTCAGCGCTCCCTGCCGAGCGGAGCGACGCTGTGTCGCT	28793
353	QY	lleuAlaLysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgTh	373
28794	Db	CGTTCGGGC-----	28808
373	QY	xlYsaGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAs	393
28809	Db	GGGAGGAGCTCGGGGGTGTCTGAGGCGCTGGGCGAGCTGTGGCGCGCGCGCTCGG	28868
393	QY	nSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArgAlaAlaValAlaMe	413
28869	Db	TCAGTGGCGCGCGCTTCCTCCACACGGTGGCGCGGTCCGCTCGCACCTATCCGT	28928
413	QY	tProProAlaSerValMetThr-----	420
28929	Db	GGAGCGGACGGATCTGGCCCGACATCGAGCTGACGCGTCGCCACGACGCGCGG	28988
421	QY	-----ArgLeuIleLeuIle-MetValTrpArgLysLeuIleArgAsn-----	434
28989	Db	ATCCGACCCAAAGCTGGTTCTATCGCTGGACTGTGGCCGGAGATACCTCGCACCTCCAGA	29048
434	QY	-----	434
29049	Db	AATCAGAGAGGAGCGCGCGGAGCTGGTGGTATTGGCGGATAAAGGTGGATGTCGCG	29108
435	QY	-----ProA	436
29109	Db	AGSCGTGCTGCAGCGCTGTCGACACGTGGACTTCCATGCTCGTCTCCATGCGCGG	29168
436	QY	snThrTyrSerSerLeuIleGlyValVal-----Tps	447
29169	Db	CAGAGACATCCCGACCCCGGAGCTGGTACCGAGGTGCGCGCGGTGCAAGCATTCGC	29228
447	QY	erLeuValSerTyrArgTrpGlyIleGluMetProAlaIleAlaArgSerIleSerI	467
29229	Db	AGTAGTGTCTACCTGTGGGTCTGGAGCGCGTGTGGTGGAGGCGTGCATCGATG	29288
467	QY	leLeuSerAspAlaGly-----LeuGlyMetAlaMetPheSerL	480
29289	Db	AGATCGCGACGCGACCCGCTCGTGTACCGCGCGGTGCTGGCTGGCTGGCTTCGTG-	29347
480	QY	eucGlyLeuPheMetAlaLeuGlnProArgIleAlaCysGlyAsnLysLeuAlaAlaI	500
29348	Db	-----AGCACGTGCTGTTCGTCGCGCGCTCGGCTCGGCTCGGCTCGGCTCG	29381

Qy 500 lealaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAAserIleAlaVal 520  
 Db 29382 TGACCCGGGGGCGATGATCGTTGGCGACGAGCTCGGATCGCCCTTGTTCAGCGCGCGT 29441  
 Qy 520 al 520  
 Db 29442 TA 29443

RESULT 7  
 US-09-568-472-1  
 ; Sequence 1, Application US/09568472  
 ; Patent No. 6358719  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Schupp, Thomas  
 ; APPLICANT: Ligon, James  
 ; APPLICANT: Molnar, Istvan  
 ; APPLICANT: Zirkle, Ross  
 ; APPLICANT: Cyr, Devon  
 ; APPLICANT: Goerlach, Joern  
 ; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
 ; FILE REFERENCE: 4-30582A  
 ; CURRENT APPLICATION NUMBER: US/09/568,472  
 ; CURRENT FILING DATE: 2000-05-10  
 ; PRIOR APPLICATION NUMBER: 09/335,409  
 ; PRIOR FILING DATE: 1999-06-17  
 ; NUMBER OF SEQ ID NOS: 30  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 68750  
 ; TYPE: DNA  
 ; ORGANISM: Sorangium cellulosum  
 US-09-568-472-1

Alignment Scores:  
 Pred. No.: 0.047 Length: 68750  
 Score: 152.50 Matches: 106  
 Percent Similarity: 29.68% Conservative: 43  
 Best Local Similarity: 21.12% Mismatches: 164  
 Query Match: 5.24% Indels: 190  
 DB: 4 Gaps: 14

US-10-030-884-14 (1-573) x US-09-568-472-1 (1-68750)

Qy 118 ValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMet 137  
 Db 28213 GTGCGTGGCGGGGTTCGCTGGAAGATGGGTGAGGCTCGTGGCGGCGC----- 28266

Qy 138 ValGlnValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuGlu 157  
 Db 28266 ----- 28266

Qy 158 TyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAAserIle 177  
 Db 28267 ---CGGGCGGTGATGAGGGCTCTCGCGGGCGGCGCGATGCTCGCTCGGAGC--- 28320

Qy 178 ValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeuGlu 197  
 Db 28320 ----- 28320

Qy 198 AlaGluProAspGlyValAlaGlyAlaGlyAlaVal----- 209  
 Db 28321 GCCGAGGCGGAGGTGGCGGGCGGTGGCGCGCACGCGCGTGGTGTGATCGCGGC 28380

Qy 210 -----SerSerArgGlyAspAlaGlyArg-----ValArg 220  
 Db 28381 GGTCAATGGCGGAGCAGGTGTGATCGCGGGCGTGGAGCAAGCGGTGGAGGGATCGC 28440

Qy 221 ValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys-SerHis---SerHi 239  
 Db 28441 GCGGGGTTTCGGCGGCGGCGGCGGACCAAGCGGTGATGCTCTCGCAGCGGTCCCA 28500

Qy 239 sSerGlnThrMetGlnPro-----ArgValSerAsnLeuSerGlyVa 253

Db 28501 CTCGCGCGTGTAGAACCGATGCTGGAGAGTTCGGGCGGTGGCGGCG-ATCGGTGACGT 28559  
 Qy 253 lGlulleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHi 273  
 Db 28560 ACCGGCGGCCAAGCGTTTCGTGTGTGAGCAACCTGAGCGGGAAGGTGTGCGC---GGACG 28616  
 Qy 273 sAlaAspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaGlyAs 293  
 Db 28617 AGCTGAGCGCGCGGGGTACTGGGTGGCGCACGTGGCGGAGCGGTGCGTTCGGCGACG 28676  
 Qy 293 pGluGlnLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGlnAlaVa 313  
 Db 28677 CGGTGAAGCGCTGCACGAAGCCGTGCGGCGACGCTTCGTCAAGTGGCGCGGACCGA 28736  
 Qy 313 lAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSerPr 333  
 Db 28737 CGCTGCT----- 28743  
 Qy 333 oValSerGluArgAlaAlaValHisValPheGlyAlaGlyAlaAspHisAlaAspVa 353  
 Db 28744 -----CGGGCTGTTGCCAGCTGCTGCCGAGCGGAGCGGACGCTGTGGCGT 28793  
 Qy 353 lLeuAlaLysGlyAlaGlnAlaTyrAspLnuTyrGlyArgAspAspTyrSerSerArgTh 373  
 Db 28794 CGTTGGCGGC----- 28808  
 Qy 373 rLysAsnGlySerGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAs 393  
 Db 28809 GCGAGAGGCTCGGGGGTGTCTCGAGGCGGTGCGGAGCGGTGCGGCGCGCGCTCGG 28868  
 Qy 393 nSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArgAlaAlaValAlaMe 413  
 Db 28869 TCAGTGGCGGGCGTCTTCCACCGCTGGCGGCGGTGCGCTGCCGTACCTCGACCTATCGT 28928  
 Qy 413 tProProAlaSerValMetThr----- 420  
 Db 28929 GCGAGCGGCGAGGTACTGGCCCGCATCGAGCGCTGACAGCGTGCACAGCGAGCGCGG 28988  
 Qy 421 -----ArgLeulleLeulle-MetValTrpArgLysLeulleArgAsn----- 434  
 Db 28989 ATCCGACCCAGGCTGTTCTATCGGTGAGTGGCGCGGAGATACCTCGCAGCCTCCAGA 29048  
 Qy 434 ----- 434  
 Db 29049 AATCAGAGGAGCGAGCCGCGGAGCTGGCTGGTATTGCGGATAGGGTGGAGTCGGCG 29108  
 Qy 435 -----ProA 436  
 Db 29109 AGCGGTGCTGCGAGCGCTGTCGACACGTGAGTCTTCATCGTCTGCTCCATGCGCGG 29168  
 Qy 436 snThrTyrSerSerLeulleGlyValVal-----TIPS 447  
 Db 29169 CAGAGCATCCGCGAGCCCGCGAGCTGGTACCGAGGCTCCCGCGGTGCGAGCGATTGC 29228  
 Qy 447 erLeuValSerTyrArgTrpGlyIleGluMetProAlaIleIleAlaArgSerIleSerI 467  
 Db 29229 AGGTAGTGTCTACCTGTGGGTGCTGGACCGCTGCTCGGTGCGGAGCGGTGATCGATG 29288  
 Qy 467 leLeuSerAspAlaGly-----LeuGlyMetAlaMetPheSerL 480  
 Db 29289 AGATCGCGCAGCGAGCCCTCGTGTACTACCGCGCGGTGCTCGCTGGTTCGTTTCG- 29347  
 Qy 480 euGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaAlaI 500  
 Db 29348 -----AGCACCGTGTCTGTTGTCGCCCGGCTCGGCTGCG 29381  
 Qy 500 lealaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAAserIleAlaV 520  
 Db 29382 TGACCCGGGGGCGATGATCGTTGGCGACGAGCTCGGATCGCCCTTGTTCAGCGCGCGT 29441  
 Qy 520 al 520  
 Db 520 al 520





```

; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 200          Length: 4403765
Score: 144.00          Matches: 121
Percent Similarity: 32.49%      Conservative: 45
Best Local Similarity: 23.68%   Mismatches: 225
Query Match: 4.95%             Indels: 120
DB: 3                        Gaps: 20

US-10-030-884-14 (1-573) x US-09-103-840A-2 (1-4403765)
QY 73 AlaAlaAspThrLeuGlnLysValAlaValLeuAlaLeuAlaLeuAlaSerArgGly 92
Db 3936616 GCCGCGGACCGCTTGGCCGCGTTCGCCAGCGGTGGCGTTCAGCCGCCACCGCGCC 3936557
QY 93 LeuSerSerProArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThr 112
Db 3936556 GCCAGCCCGCGCGCCACCGGC-----ACCACCGCTCGCGCTTGACCTCC 3936509
QY 113 LeuProAsnThrLeuValMetGlyIleProLeuLeuArgGlyMetTyrglyAlaSerSer 132
Db 3936508 TTGCCC---GCCGCTGCCGCGGACGTGGT-----GGCCCGCTGCC 3936467
QY 133 AlaGlyThrLeuMetValGlnValValLeuGlnCysIleIleTrpTyThrLeuMet 152
Db 3936466 GCCGCGCGCGCCGACCGCGAC----- 3936443
QY 153 LeuPheLeuPheGluTyArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGly 172
Db 3936442 -----TTGCGCGTCCGCGCGTTCGCGCGCTTCCCGCCGCGACCGCATACC 3936392
QY 173 AlaAlaAlaSerIleValSerPheArgValAspSerValValSerLeuAlaArgGly 192
Db 3936391 GTCCGCGCGCGC-----GCCACCACTCGCGCTTGC CGCGGT 3936353
QY 193 AspValGluLeuGluAlaGluProAspGlyValAlaGlyAlaGly-----AlaValSer 210
Db 3936352 -----GCCGCGGTTCGCGCGCGCGCGCGCGTTCGCGCGCGCGCGCGCGTTCGTCG 3936305
QY 211 SerArgGlyGlyAsp-----AlaGlyArgVal----- 219
Db 3936304 AGCCGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3936245
QY 220 -----ArgValThrValArgLysSerThrSerArgSerGluAlaAlaCysSer 236
Db 3936244 GACGTACCGCGGTTCGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3936191
QY 237 HisSerHisSerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTy 256
Db 3936190 ---TCCGCAGTCCACCGCTTCCGCGCGCTATTTCGCTACCAACCGCGCGTCTCGCGCAGC 3936134

```

```

257 SerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhe 276
3936133 GCCCGCATGGCGCGGTTCGCGCGGTTCGCGCGGTACCTCCGTCGCGCGCGCGCGCC 3936074
277 PheAsnIleValGlyAlaAlaLysGlyGlyGlyAlaAlaGlyAspGlu----- 294
3936073 GTGATGCCGTACCGCGCAGCAGCGCGGTTCACCGGTTCGCGCGCGCGCGCGCGGTTC 3936014
295 -----GluLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 309
3936013 GCCATGCCCGCGCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTC 3935954
310 ProGlnAlaValAla-----ValProAlaLysArgLysAspLeuHis 323
3935953 ATCTCCGCGCGCGCGCGCATTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTC 3935894
324 MetLeuValTrpSerSerSerAlaSerProValSerLeuArgAlaAlaValHisValPhe 343
3935893 GCCCGCGCGCGCGCGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCG 3935834
344 GlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTyArgGlu 363
3935833 GGTGCTGCTGCTGACGCGCTGCTCCACCGCTTCGCGCGGTTCGCGCGGTTCGCGCGGTTC 3935783
364 TyrGlyArgAspAspTySerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGly 383
3935782 -----GCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3935747
384 GlyPro-----ThrLeuSerLysLeuGlySerAsnSerThr----- 395
3935746 GCGCGCGCGCTACCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTC 3935687
396 AlaGlnLeuTyProLysAsp-----AspGlyGluGlyArgAlaAlaAlaValAlaMetPro 414
3935686 GCCGCTGCTACCGCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTC 3935627
415 ProAlaSerValMetThrArgLeuLeuLeuMetValTrpArgLysLeuIleArgAsn 434
3935626 CCATCGCGCGCGCGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCG 3935597
435 ProAsnThrTySerSerLeuIleGlyValValTrpSerLeuValSerTyArgTrpGly 454
3935596 AGCGAGCGCGTACCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGT 3935540
455 -----IleGluMetProAla 459
3935539 ACCACCGTGTGCTGCGTGCAGTGCACGTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3935480
460 IleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSer 479
3935479 GCCGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCG 3935420
480 LeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCys----- 493
3935419 GCCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3935360
494 GlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMet 513
3935359 GGTGGTGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCG 3935300
514 AlaAlaAlaSerIleAlaValGlyLeuArgGly 524
3935299 GCCGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGGTTCGCG 3935267

```

## RESULT 11

```

US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.

```

**iii**

**iii**



1151	CAAAATGAGTGGTATAAAGGACCGCATTTAGGAGACATTAAGACCTTCTCAGGAGTAATGGA	1211
319	-----	319
1211	GCCTCTGCAAAAGCCTCTGCTAAAGCATCAGCAGTAGCAAGCACAAAATCTCAAATTGAC	1270
320	-----	320
1271	GATTTGAAGATGCTTTAAGAGATCTTGCAAGTCTATTGAAGAGCTCAGCACTGCTTCA	1330
320	-----	320
1331	GCATCTGCATCTGCATCAGCTTCAGCTGAGGTGGAGCGGTGGTGTAAACGCGAGTGGT	1390
351	-----	351
1391	AACGGAGGAGGCGCGGTGGAGCTGAGCTCTAGCTGCTCTCGCTGCTCGCAGGA	1450
363	GlutyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLys	382
1451	GCCTGAGGT-----GGACTTGAGGTGGAGCGCGGA	1480
383	GlyGlyProThrLeuSerLysLeu-----GlySerAsnSer	394
1481	GGCGGAGCTTTAGCCCGCTGCATAGCTGCTGCTGTCAGGTGGAGGAGGTTTGGTGGGA	1540
395	ThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetPro	414
1541	CTTGGAGGACTAGCGCGTCTTGGTGGGGATCTGCCCGAGCTGCTGCACCGCTGCGCT	1600
415	ProAlaSerValMetThrArgLeulleLeulleMetValTrpArgLysLeulleArgAsn	434
1601	GCTGCATCAGGTGGTGGAGGAGCAGCTTAGAAGGCTTTGAGAAGACAAATGCGTGGGA	1660
435	ProAsnThrTyrSerSerLeulleGlyValValTrpSerLeuValSerTyrArgTrpGly	454
1661	GGTGGATCGCGCTGCTCCCGCTGCTGCTGCTGTCAGCTGCTGTCGAGGTGGATGGGA	1720
455	IleGluMetProAlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGly	474
1721	GGTGGAAATG-----GGTGGAGGATTCGGA	1744
475	MetAlaMet-----PheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIle	491
1745	GTAGGTCTCGTGGAGGATTCGGAGGAGGATTT-----	1777
492	AlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheValAlaGly	509
1778	---GGTGGTGGATCATCAGCAGCAGCTGCTGCCGTGCTGTCAGCCCGCTGGATTGGT	1834
509	-----	509
1835	GGAGGTGGACGAAGAGGTAGAGGTAGAGGACGTGGAGCGATGGCGACGTAAACGAGCT	1894
510	ProAlaValMetAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIle	529
1895	AGTGTCTAGCTGCAGCCCGCGCTGCTGCTGCTGCTGCGAGGATCTGCTGCTGATGT	1954
530	AlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGluTyr	549
1955	GCCTGCGCTGCTGCTGCACCC-----GCAGCTATGTATC	1987
550	GlyValHisProAspIleLeuSerThrAlaTyrGlyProIleThrSerHisGlyPhe	568
1988	GSTGACGGTGTGAT-----GGACCTGATTTTCGATATGATTC	2026

RESULT 14  
 US-08-864-038A-4  
 ; Sequence 4, Application US/08864038A  
 ; Patent No. 6001592  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kunio NAKASHIMA et al.





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 3, 2004; 11:11:45; Search time 1319 Seconds  
(without alignments)  
1588.862 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2310

Sequence: 1 MIVALDIYHVLTVVPLVYA.....DILSTAYGPTSHGFITCHS 573

Scoring table:

BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Rgapop 6.0, Rgapext 7.0

Delop 6.0, Delext 7.0

Searched: 2421054 seqs, 1828716029 residues

Total number of hits satisfying chosen parameters: 4842108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+\_p2n.model -DEV=xlh

-Q=/cgn2\_1/USPTO\_spool/US10030884/runat\_24022004\_103926\_2050/app\_query.fasta\_1.711

-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1

-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0

-MAXLEN=200000000 -USER=US10030884@cgn1.1.57 @runat\_24022004\_103926\_2050

-NCPU=6 -ICPU=3 -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSBLOCK=100

-LOGLOG -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*

2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*

3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*

4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*

5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*

6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*

7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*

8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*

9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*

10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*

11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*

12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*

13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*

14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*

15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*

16: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*

17: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

18: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
-----					

1	2129.5	73.2	1926	15	US-10-260-238-565	Sequence 555, App
2	1822	62.6	2346	12	US-10-424-599-134028	Sequence 134028, A
3	1643	56.5	3020	12	US-10-424-599-45725	Sequence 45725, A
4	1621.5	55.7	3374	12	US-10-424-599-86575	Sequence 86575, A
5	1603.5	55.1	1759	15	US-10-260-238-30	Sequence 30, App1
6	1591.5	54.7	2415	12	US-10-425-114-16771	Sequence 16771, A
7	1560.5	53.6	1923	9	US-09-938-842A-847	Sequence 847, App
8	1542.5	53.6	1923	11	US-09-938-842A-847	Sequence 847, App
9	1542.5	53.0	2272	9	US-09-887-576-628	Sequence 628, App
10	1542.5	53.0	2272	12	US-10-425-114-1349	Sequence 1349, App
11	1536	52.4	1860	9	US-09-938-842A-848	Sequence 848, App
12	1536	52.4	1860	11	US-09-938-842A-848	Sequence 848, App
13	1406.5	48.3	1983	9	US-09-938-842A-1305	Sequence 1305, App
14	1406.5	48.3	1983	11	US-09-938-842A-1305	Sequence 1305, App
15	1272.5	43.7	1928	12	US-10-425-114-10077	Sequence 10077, A
16	1147	39.4	1744	12	US-10-425-114-7737	Sequence 7737, App
17	1027	35.3	1273	12	US-10-425-114-15739	Sequence 15739, A
18	1004.5	34.5	1334	12	US-10-425-114-2319	Sequence 2319, App
19	1004	34.5	1321	12	US-10-425-114-33502	Sequence 33502, A
20	1003	34.5	1199	12	US-10-425-114-29155	Sequence 29155, A
21	1002.5	34.5	1313	12	US-10-425-114-1838	Sequence 1838, A
22	1000	34.4	1091	12	US-10-425-114-3558	Sequence 3558, App
23	987	33.9	1651	15	US-10-260-238-5337	Sequence 5337, App
24	946.5	32.5	1539	12	US-10-425-114-29226	Sequence 29226, A
25	937.5	32.2	1464	12	US-10-424-599-104351	Sequence 104351, A
26	892.5	30.7	1146	12	US-10-424-599-59026	Sequence 59026, A
27	825	28.4	1271	12	US-10-424-599-18686	Sequence 18686, A
28	795.5	27.3	736	12	US-10-424-599-134033	Sequence 134033, A
29	772.5	26.5	1280	12	US-10-424-599-45727	Sequence 45727, A
30	767.5	26.4	1122	12	US-10-424-599-134029	Sequence 134029, A
31	757	26.0	702	15	US-10-260-238-5464	Sequence 5464, App
32	733.5	25.2	1211	12	US-10-424-599-45726	Sequence 45726, A
33	611	21.0	971	12	US-10-424-599-62375	Sequence 62375, A
34	590	20.3	1045	15	US-10-260-238-4000	Sequence 4000, App
35	568.5	19.5	761	12	US-10-425-114-31117	Sequence 31117, A
36	561	19.3	987	12	US-10-425-114-13089	Sequence 13089, A
37	521	17.9	729	12	US-10-424-599-59027	Sequence 59027, A
38	513	17.6	673	12	US-10-424-599-71900	Sequence 71900, A
39	465	16.4	606	12	US-10-424-599-34427	Sequence 34427, A
40	447	15.0	558	12	US-10-424-599-52693	Sequence 52693, A
41	427.5	14.7	650	12	US-10-424-599-33472	Sequence 33472, A
42	410	14.1	869	12	US-10-424-599-84904	Sequence 84904, A
43	365	12.5	644	12	US-10-425-114-30762	Sequence 30762, A
44	328	11.3	574	15	US-10-260-238-5138	Sequence 5138, App
45	323	11.1	2000	15	US-10-260-238-2559	Sequence 2559, App

#### ALIGNMENTS

##### RESULT 1

US-10-260-238-565

; Sequence 565, Application US/10260238

; Publication No. US20040016025A1

; GENERAL INFORMATION:

; APPLICANT: Budworth, Paul R.

; APPLICANT: Moughamer, Todd G.

; APPLICANT: Briggs, Steven P.

; APPLICANT: Cooper, Bret

; APPLICANT: Glazebrook, Jane

; APPLICANT: Goff, Stephen A.

; APPLICANT: Katagiri, Fumiyaki

; APPLICANT: Krepis, Joel

; APPLICANT: Provart, Nicholas

; APPLICANT: Ricke, Darrell

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION

; FILE REFERENCE: 60111-NP

; CURRENT APPLICATION NUMBER: US/10/260,238

; CURRENT FILING DATE: 2002-09-26

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US 60/325,448

; PRIOR FILING DATE: 2001-09-26

```
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 565
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (261)..(261)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (291)..(291)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (486)..(486)
; OTHER INFORMATION: n = any nucleotide
US-10-260-238-565

Alignment Scores:
Pred. No.: 1,4e-211 Length: 1926
Score: 2129.50 Matches: 458
Percent Similarity: 77.58% Conservative: 23
Best Local Similarity: 73.87% Mismatches: 42
Query Match: 73.18% Indels: 97
DB: 15 Gaps: 16

US-10-030-884-14 (1-573) x US-10-260-238-565 (1-1926)
QY 1 MetIleThAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
DB 1 ATGATCACGGTGGTGACCTGTACACAGCTCTGACGGCGGTGGTGGTGTACGTGGG 60
QY 21 MetThLeuAlaTyrGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40
DB 61 ATGACGCTGGCGTACGCTCCGTCCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTCC 120
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
DB 121 GGATCAACCGCTTCGTGGCGCTTCCTGCGCTCCCGCTCTCTCTCTCTCTCTCTCT 180
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
DB 181 ACCAACHACCCCTTCGCCATGAACCTCCGCTCCTCGCGCGGACACGCTCCAGAGCTC 240
QY 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96
DB 241 ATGCTCTCTGCGCTCTCTGCGCTCTGCGCTCTCTGCGCGCGG-----288
QY 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116
DB 289 -----TCNCTGACTGGCTCATCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 339
QY 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAla-----133
DB 340 CTCGTATGGGACATCCGCTGTCTCAAGGGAGATGACGCGCGCGCGCGCGCGCGCGCG 399
QY 134 -----GlyThrLeuMetValGlnValValLeuGlnCysIleIleTrpTyrThr 150
DB 400 GCCGACTCCGCGACCTCATGGTGCAGATCGTGTCTCCAGTGCATCATCTGTGTACAG 459
QY 151 LeuMetLeuPheLeuPheGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPhePro 170
DB 460 CTCATGCTGTCTCTTCGAGTACCGNGCGCGCGCTCTCTGTCTCATGAGCAGATTCCCG 519
QY 171 AspGlyAlaAlaAlaSerIleValSerPheArgValAspSerAspValValSerLeuAla 190
DB 520 GAC-----ACCGCGCGCTCCATCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 576
QY 191 ArgGlyAsp-----ValGluLeuGluAlaGluProAspGlyValAlaGlyAlaGly 207
```

```
577 GCGCGCGGTGGTGAGCCCGGAGCTGCAGCGCGAG-----612
208 AlaValSerSerArgGlyGlyAspAlaGlyArgValArgValThrValArgLysSerThr 227
613 -----GCGGAGGTGGGAGCAGATGCAGATCGGGTCCACCGTGCAGTGCAGTGCAG 663
228 SerSerArgSerGluAlaAlaCysSerHis-----SerHisSerGlnThrMetGln 244
664 AGTTCGCGCTCCGAGCGCGGTCTCGACGGACGACGTCACACTCGCAGTCCATGCAG 723
245 ProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnPro 264
724 CCGCGCGTCTCCAACTCTCCGCGCTCGAGATTTACTCGTCGAGTCTCGCGGAACCCG 783
265 ThrProArgGlySerSerPheAsnHisAlaAspPhePheAsnIleValGlyAlaAlaAla 284
784 ACGCGCGTGGTCCAGCTTCAACCCAGCGAGTCTTCAACATCTCTC-----831
285 LysGlyGlyGlyGlyAlaAlaGlyAspGluGluLysGlyAlaCysGlyGlyGlyGly 304
832 -----GGCAACGGCAAGCAGCGCGAGGAGAGAGCGGCC-----GCCGCGTGGCGG 879
305 GlyHisSerProGlnProGlnAlaValAlaValProAlaLysArgLysAspLeuHisMet 324
880 GGCACACTCGCGCAGCGC-----GTGGTGGGAGAGAGAGAGACCTGCACATG 927
325 LeuValTrpSerSerAlaSerProValSerGluArgAla-----338
928 TTCGTGTGGAGCTCAAGCGCTCGCGGTGTGGAGCGCGCGCGCGCGCAGCAGCTGGC 987
339 AlaValHisValPheGlyAlaGlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAla 358
988 GCGGTGCACGCTTTCGCGGTGGTGGCGCGCAGCACCGCGAC-----GCCAAGAGAGCT 1041
359 GluAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArgThrLysAsnGlySerGly 378
1042 CAGCCCTATGATGAG-----TACAGTTTCGGGACACAGAAC-----1077
379 GlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeu 398
1078 -----GAGAAGACGCGCGCGCTGTCCAGCTGGGTGCCAAGTGCAGCGCGCAGCTC 1131
399 TyrProLysAspAspGlyGlyArgAlaAlaAlaValAlaMetProProAlaSerVal 418
1132 CGGCCAAGACGACGCGCGAGGAGTGGCGCG-----GCGATGCGCGCGCGAGCGGTG 1185
419 MetThrArgLeuIleLeuIleMetValTrpArgLysLeuIleArgAsnProAsnThrTyr 438
1186 ATGACGAGGTCTCTCTGATCATGTTTGGAGGAAGCTGATCAGGAACCCCAACACTTAC 1245
439 SerSerLeuIleGlyValValTrpSerLeuValSerTyrArgTrpGlyIleGluMetPro 458
1246 TCCAGCTCTCTGTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1305
459 AlaIleIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPhe 478
1306 GCCATCATCGCGCTCGATTTCGATCTCTCAGATGTCAGGGTTGGAATGGCCATGTC 1365
479 SerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAla 498
1366 AGCTTCGAGATTGTTCATGGCATTCAGCCAGCATCATTCCTCTGTGGGAATCTCCTTGT 1425
499 AlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIle 518
1426 TCGTATGCCATGGCGCTCAGTTCCTCTGTGGTCCGCGCTCATGTGCTGCCGCTCCATC 1485
519 AlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleVal-----532
1486 GCCGTCCGACTTCGCGGGTGTCTCTGCATTTGCCATTGTTCAGGTAAAGCTGCAAGTTC 1545
532 -----532
1546 AGGCTTCTGCACCAAAATGCATCTTAACTGCAGAAAGAAATCTATCTGTAAGTCAAAATTG 1605
```

QY 533 -----GlnAlaLeuProGlnGlyIleValProPheValPheAlaLysGlu 548  
 Db 1606 TCATTGAATTCAGGCGCTCTCCCAAGAACTCGCCCTTGTGTGTCAGAGGAG 1665  
 QY 549 TyrGlyValHisProAspIleLeuSerThrAlaTyrGlyProIleThrSerHisGlyPhe 568  
 Db 1666 TACAATGTTTCACTCAACATTCAGCAGCGATGATCTCTTTTACCATCTTTTCTTTT 1725

## RESULT 2

US-10-424-599-134028  
 ; Sequence 134028, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 134028  
 ; LENGTH: 2346  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_92032C.1  
 US-10-424-599-134028

## Alignment Scores:

Pred. No.: 1,98e-179 Length: 2346  
 Score: 1822.00 Matches: 387  
 Percent Similarity: 70.83% Conservative: 55  
 Best Local Similarity: 62.02% Mismatches: 86  
 Query Match: 62.61% Indels: 96  
 Ds: 12 Gaps: 16

US-10-030-884-14 (1-573) x US-10-424-599-134028 (1-2346)

QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
 Db 148 ATGATAACGTTGTTGGACTTGTACACCTTCTGACAGCGGTGGTCCGCTCTACGTGGCG 207  
 QY 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40  
 Db 208 ATGATCCTCGCTACGGCTCGTGAAGTGGTGAAGATCTTCAACCCCGGACCAATGCTCG 267  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 268 GGAATAAACCGGTTCTGGGCATCTTCGACGTGCCCCCTCTCTCTCCACTTCATCTCC 327  
 QY 61 ThrAsnAppProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 328 ACCAACACCCCTAGCCCTGAACCTACAGTTTCATCGCGCGGACACACTCCAAAAAGCC 387  
 QY 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
 Db 388 ATAGTCTCGAGTGTCTTGTGTGTGTGTCAGCAGC-----ACAGCTCAAGAGGC---TCA 438  
 QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120  
 Db 439 CTGTAATGTTCCATACACTCTCTCTCTCAACTCTCCCAACACCTTGGTGATGGGT 498  
 QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
 Db 499 ATCCCGTCTCGAAGGGCATGTACGGTGATGACTCG---GGAACTCAATGTTTCAGATT 555  
 QY 141 ValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPheGluTyrArgAla 160  
 Db 556 GTGGTCTTCAGTGTATCATCTGGTATACCTTATGTTGTTCTTTTCAGTATAGGGGT 615

QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
 Db 616 GCGAGGCTTTAATAGTGGAGCAGTTTCTGTAT---ACAGCGGTTCCGATTATCTCTTC 672  
 QY 181 ArgValAspSerAspValValSerLeu---AlaArgGlyAspValGluLeuGluAlaGlu 199  
 Db 673 AAGTTGATTCGATATATATTTTCGTTGATGGGAAGAGCCCTTCAGACGAGGCGAG 732  
 QY 200 ProAspGlyValAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal 219  
 Db 733 GTT-----GGTGATACGCGAAGCTT 753  
 QY 220 ArgValThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSerHis 239  
 Db 754 CATGTACCGTGAGGAAGTCGCGAGTTCGCGTTCGAGATTTTCTTAGGCGCTCTCAT 813  
 QY 240 -----SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyr 256  
 Db 814 GSCCCCAACTCGGTGTCGTGATTCGAGGCTTCGAATTTACCAATCGGAGATTAC 873  
 QY 257 SerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhe 276  
 Db 874 TCGCTCAGAGCTCGAGGAACCGACTCGAGAGGCTCGAGTTTAAACCACGGAATTT 933  
 QY 277 PheAsnIleVal----- 280  
 Db 934 TACTCCATGTTGATGTTAAGTAACAACAACAACAACAACAGTATGATGATCG 993  
 QY 281 -----GlyAlaAlaAlaLysGly 286  
 Db 994 AGGCAGAGTAATTTTGGAGGCTTTGATGAAGAAAGTGTGGGGGTGTGAGGTTGAATGT 1053  
 QY 287 GlyGlyGlyAlaAlaGly----- 292  
 Db 1054 GGTGCTGGTGTACGTGGGTACCTGGGCTCGAATCGGGGATTTTTCCTCGGTGGCG 1113  
 QY 293 AspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnProGlnAla 312  
 Db 1114 GCGAAGAAGAAGGTTGTCGAGAGTGTGTGCGAGG----- 1149  
 QY 313 ValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAlaSer 332  
 Db 1150 -----AAGGATCTCCACATGTTTGTGTGGAGTTCGAGTGTCTTCG 1188  
 QY 333 ProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAlaAspHis---Ala 351  
 Db 1189 CCGGTGTCGAA---GGTGAATCCATGCTTCAGAGGTGTGTGGGATTTATGGAGT 1245  
 QY 352 AspValLeuAlaLysGlyAlaGlnAla-----TyrAspGluTyrGlyArgAsp 367  
 Db 1246 GACCAGCTTCTGTTGGTGGGTGCTCACCAAGAAAGATTATGATGATGTTGTTCAGAT 1305  
 QY 368 AspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeu 387  
 Db 1306 GAGTTTAGCTTCGGGAACAGAACCGTTGCTAATGGGTTGACAAAGAGGCGGAGTCTT 1365  
 QY 388 SerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGlyArg 407  
 Db 1366 TCAAGGCTTGGCTCGAGTTCACAGCTGAGCTTCGCCCTTAAAGCTCAAGGTGAA----- 1419  
 QY 408 AlaAlaAlaValAlaMetProProAlaSerValMetThrArgLeuIleLeuMetVal 427  
 Db 1420 GCCAAACCTACTTCCATGCCACCAACAGTGTATGACAGGCTCATTTTGTATGATT 1479  
 QY 428 TrpArgLysLeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSer 447  
 Db 1480 TGGAGGAAGCTGATTAGGAACCCCAACACATATTCAGGCTCTTTGCTCTCACTTGGTCT 1539  
 QY 448 LeuValSerTyrArgTrpGlyIleGluMetProAlaIleIleAlaArgSerIleSerIle 467  
 Db 1540 TTGATCTCAATCAATGAATGTTGTATGCCAGCAATGTTGCTAAATCGATATCAATT 1599

QY 468 LeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGln 487  
 Db 1600 TTATCTGATGCTGGTCTTGGGATGGCAATGTTTAGCCTTGGGTATTATTCATGGCAATTGGAG 1659  
 QY 488 ProArgIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheVal 507  
 Db 1660 CCAAGATTATTCATGTGCAAGTGAACCTCGGTGCTTCTTGTCTATGGCAGTTGTTTCTT 1719  
 QY 508 AlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeu 527  
 Db 1720 ACTGGTCTCGCATGCTATGCTGCTGCTTCAATCGTTAGGGCTCAGGGGAGTTCTATTG 1779  
 QY 528 HisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLys 547  
 Db 1780 CACATTGCTATTGTACAGGCTGCTGCTCAGGGATGTCCTTTTGTGTTGCTTAAG 1839  
 QY 548 GluTyrGlyValHisProAspIleLeuSerThrAla-----TyrGly----- 561  
 Db 1840 GAATACAGGTTTCATCTGACATACCTAGACCGGGGTTATATTGGGATGCTAATTGCT 1899  
 QY 562 ---ProIleThr 564  
 Db 1900 CTTCTATTACG 1911

RESULT 3  
 US-10-424-599-45725  
 ; Sequence 45725, Application US/10424599  
 ; Publication No. US20040031072A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa Thomas J  
 ; APPLICANT: Kovalic David K  
 ; APPLICANT: Zhou Yihua  
 ; APPLICANT: Cao Yongwei  
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53223)B  
 ; CURRENT APPLICATION NUMBER: US/10/424,599  
 ; CURRENT FILING DATE: 2003-04-28  
 ; NUMBER OF SEQ ID NOS: 285684  
 ; SEQ ID NO 45725  
 ; LENGTH: 3020  
 ; TYPE: DNA  
 ; ORGANISM: Glycine max  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_141290C.1  
 US-10-424-599-45725

Alignment Scores:  
 Pred. No.: 1-26e-160 Length: 3020  
 Score: 1643.00 Matches: 361  
 Percent Similarity: 66.51% Conservative: 58  
 Best Local Similarity: 57.30% Mismatches: 111  
 Query Match: 56.46% Indels: 100  
 DB: 12 Gaps: 17

US-10-030-884-14 (1-573) x US-10-424-599-45725 (1-3020)

QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20  
 Db 755 ATGATCACCTTAACAGACTTCAACCATGTCATGACATGCAATGGTGCCTATGTGGCC 814  
 QY 21 MetThrLeuAlaTyrGlySerValArgTTPTrpArgIlePheThrProAspGlnCysSer 40  
 Db 815 ATGATACTAGCTATGGCTCAGTGAAGTGGTGGAGATTTCCTCCCTGACCAATGCTCT 874  
 QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
 Db 875 GGCATCAACCGTTTGTGGCACTCTTTGCACTGCTCTCTCTCTCCCTTCACCTTCATAGCC 934  
 QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
 Db 935 TCCAAACACCCCTATGAGATGAACCTGAGGTTCTTAGCTGCTGACACCCCTTCAAAGATC 994

QY 81 AlaValLeuAlaLeuLeuAla-----LeuAlaSerArgGlyLeuSerSerPro 96  
 Db 995 ATAATACTAGTCTCTCTTGGCAGTTGGAGCAACATCACCAAAAGGGGTGT----- 1045  
 QY 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
 Db 1046 -----TTGGAATGGGCCATAAACCCTTCTCTCTCTCCACCCCTCCCAACACT 1093  
 QY 117 LeuValMetClyIleProLeuLeuArgGlyMetTyrClyAlaSerSerAlaGlyThrLeu 136  
 Db 1094 TTGGTTATGGGCATCCCTTTGCTCAAAGGATGATATGGTGACTTCTCA-----GGAGCCTC 1150  
 QY 137 MetValGlnValValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156  
 Db 1151 ATGGTGCAAAATTGGTCTCCAGTGCATCATTTGGTACACCTTGATGCTCTCTTCTTGT 1210  
 QY 157 GluTyrArgAlaAlaAlaValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
 Db 1211 GAGTTAGAGGTGCCGAATGCTCATCTCTGACGAGTTCCTCTGAC-----ACTGCTGCTCTC 1267  
 QY 177 IleValSerPheArgValAspSerAspValValSerLeu-----AlaArgGlyAspValGlu 195  
 Db 1268 ATTGCTCCATCCATGTGGACTCTGATGTCATGTCATTTGGATGGAAGACCAACCTTGAG 1327  
 QY 196 LeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAsp 215  
 Db 1328 ACTGAAGCTCAGATCAAG-----GAA 1348  
 QY 216 AlaGlyArgValArgValThrValArgLysSerThrSerArgSerGluAlaAlaCys 235  
 Db 1349 GATGGTAACCTCCATGTCATCTGTGAGAAATCAATGCTTCAGATCAGACATCTTCTCT 1408  
 QY 236 SerHisSerHisSer---GlnThrMetGlnProArgValSerAsnLeuSerGlyValGlu 254  
 Db 1409 AGAAGGTCTCAGGCTCTCTCTCCACCACTCCACGCGCTTCCAAACCTTACCATGCTGAG 1468  
 QY 255 IleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAla 274  
 Db 1469 ATATACTCTTTGCAATCTCTAGAAACCTTACCCGAGAGGCTCTAGTTTCACACCACT 1528  
 QY 275 AspPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyGlyAlaAlaGlyAsp--- 293  
 Db 1529 GAATTTACTCCTCATGATGGCTGCTGGTGGCAGGAACCTTGGTCTCTGATGTT 1588  
 QY 294 -----GluGluLysGly 297  
 Db 1589 TATGGCCTTTTCAAGAGGGCCAACTCCAAAGGCTTTTCAACTATGATGAAGATGTT 1648  
 QY 298 Ala-----CysGlyGlyGlyGlyGlyHisSerProGlnProGln 311  
 Db 1649 GGGAGGCCAAAGTTTCATTACCATGCTGCTGTGTGGAACCTGGGCACTACCTGACCTAAC 1708  
 QY 312 -----AlaValAlaValProAlaLys 318  
 Db 1709 CTTGGCATGTTCTCTCCCTCTAATGGGTCCAAAGTGTGCTCTAATGCTAATGCCAAG 1768  
 QY 319 Arg-----LysAspLeuHisMetLeu 325  
 Db 1769 AGGCCTAATGGCAGGCTCAGCTGAAGCCTGAGGATAGGGAATAGGCACTTCATATGTTT 1828  
 QY 326 ValTrpSerSerAlaSerProValSerGluArgAlaAlaValHisValPheGlyAla 345  
 Db 1829 GTTTGGAGTTCAAGTGTCTTCCAGCTCTCTGACGCTGTTGGTCCCTCATGATGATGA--- 1885  
 QY 346 GlyGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAlaTyrAspGluTyrGly 365  
 Db 1886 GGAGTTCATGATCAGAAAGAGTCAATGATGATATCTCCAGGAAAAGTGGAGATAAT 1945  
 QY 366 -----ArgAspAspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAsp 381  
 Db 1946 CATAGGACACTCAAGAAAGACTACCTAGAGAAAGATGAGTTTCAGCTTTGGGAATAGAGAA 2005  
 QY 382 LysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLys 401

```

Db      2006  ATG-----GACAGGAGAGATGATCAGCTTGAAGGTGAG  2038
Qy      402  AspAspGlyGluValArgAlaAlaValAlaMetProProAlaSerValMetThrArg  421
Db      2039  AAGTTGGAGATGGAAACCAAA-----ACCATGGCTCCAGCAAGTGATGACAGG  2092
Qy      422  LeuIleLeuIleMetValTTPArgLysLeuIleArgAsnProIleThrTyrSerSerLeu  441
Db      2093  CTTATATTGATTATGGTGTGGAGAAACTCATCAGAAACCCCAACACTACTCTAGCCTA  2152
Qy      442  IleGlyValValTTPSerLeuValSerTyrArgTTPGlyIleGluMetProAlaIleIle  461
Db      2153  ATTGGTCTCACTTGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT  2212
Qy      462  AlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGly  481
Db      2213  GCAAAAGTCTATCTCCATATTGTCAGACGAGGGTGGCATGGCCATGTTTCAGTCTTGGT  2272
Qy      482  LeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAla  501
Db      2273  CTCCTTCATGGCTTGGCAACCGAGGTGATGATGATGATGATGATGATGATGATGATGAT  2332
Qy      502  MetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGly  521
Db      2333  ATGGCTGTGAGATTCCTTACAGGTCAGGTCATGTCATGTCATGTCATGTCATGTCATGTC  2392
Qy      522  LeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleVal  541
Db      2393  CTCAAAGTGTTCTCCATACAGTTGCCATTGTTCCAGGAGCTCTTCCCAAGGAATTGTC  2452
Qy      542  ProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAla-----  559
Db      2453  CCATTGTCTTGTGTAGGAATATAATGATATCTGATCTGATATCTCAGCAGAGCTGTATT  2512
Qy      560  TyrGly-----ProIleThr  564
Db      2513  TTTGGATGCTGATTGCTTTTCCCAATAACT  2542

RESULT 4
US-10-424-599-86575
; Sequence 86575, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ. ID NOS: 285684
; SEQ ID NO 86575
; LENGTH: 3374
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_49189C.1
US-10-424-599-86575

Alignment Scores:
Pred. No.:      2,56e-158      Length:      3374
Score:          1621.50      Matches:      375
Percent Similarity: 62.68%      Conservative: 60
Best Local Similarity: 54.03%      Mismatches: 94
Query Match:      55.72%      Indels:      165
DB:              12          Gaps:          21

US-10-030-884-14 (1-573) x US-10-424-599-86575 (1-3374)

Qy      1  MetIleThrAlaLeuAspLeuTyrHis-ValLeuThrAlaValProLeuTyrValAl  20
```

```

Db      335  ATGTAACCTGGGGAGATTTCTATCACGGTGTATGTCGGGTGATTCGGCTGTACGTGGC  394
Qy      20  aMetThrLeuAlaTyrGlySerValArgTTPArgIlePheThrProAspGlnCysSe  40
Db      395  GATGATCTTAGGTAAGGCTCGGTGGGTGGAGATATTCTCGCGGACCAATGCTC  454
Qy      40  rGlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSe  60
Db      455  CGGCATAAACCCGCTTCGTCGGGATCTTCGGCTTCGCTCTCTTCTGTTCACTTCATCTC  514
Qy      60  rThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVa  80
Db      515  CTTGAAACATCCCTACGCGATGAACCTCGGTTTCATCGCTCGGACACGCTCCAGAAAT  574
Qy      80  lAlaValLeuAlaLeuLeuAlaLeu-----AlaSerArgGlyLeuSerSerPr  96
Db      575  CATCATGCTCTTCGGCCCTCGCAATCTGGACGAACCTCTCCGCGCAACGGG-----  623
Qy      96  oArgAlaLeuGlyLeuAspTTPSerIleThrLeuPheSerLeuSerThrLeuProAsnTh  116
Db      624  -----ACCTGGAGTGAATGATTTACCAATTTTCTCCCTCTCCATTCGCGGACAC  673
Qy      116  rLeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrIle  136
Db      674  GCTCGTGAATGGGAATCCCGTGTCTCATCGCCATGTACGCGAATACTCC---GGCTTATT  730
Qy      136  uMetValGlnValValValLeuGlnCysIleIleTTPThrLeuMetLeuPheLeuPh  156
Db      731  GATGTGCAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT  790
Qy      156  eGluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSe  176
Db      791  CGAGTACCGCGCGCGGAAGCTTCTTATAATGGAACAGTTTCTCTGAA---ACCGCTGCTC  847
Qy      176  rIleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGl  195
Db      848  TATAGTCTCTTTCAAGTTGATTCGAGCTCGTTTCGCTCGATGGAGAGGATTTTCTCGA  907
Qy      195  uLeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyLys  215
Db      908  AACGGATGCCGAAGTC-----GGCGA  928
Qy      215  pAlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCy  235
Db      929  TGATGGGAAGCTTCACGTGACCGGTTAGGAAGTCGAATGCATCTAGA-----  974
Qy      235  sSerHisSerHisSerGlnThrMetGlnProArgValSerLeuSerGlyValGluIle  255
Db      975  -----AGTCTCTTTATGATCACTCCGCGGCCCTCTTAACCTCACCGCGCGGAGAT  1024
Qy      255  eTyrSerLeuGlnSerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAs  275
Db      1025  TTAACAGCTCAGCTCGTCGGTAAACCCACGCGCACGTGGCTCCAACCTTCAACCCAGCG  1084
Qy      275  pPhePheAsnIleValGlyAlaAlaAlaLysGlyGlyGlyAlaAlaGlyAsp-----  293
Db      1085  TTTTACTCCATGATGGGGGTACGCGCGCGGATCTTAATTTTGTGGCGCGGATTGTA  1144
Qy      294  -----GluGluLysGl  297
Db      1145  CTCGGTGCAGTCCAGTCACGCGGGGTACACCGCGCCCTCGAACTTCGAGGAGAACGG  1204
Qy      297  yAla-----  298
Db      1205  TGGCGCTGTGCTGCAACCAACGACGACGACGAGGCCATCAGCTCCCGAGGTTTCG  1264
Qy      299  -----CysGlyGlyGlyGlyGlyHis-----  306
Db      1265  GGTTCATCCCGCGGTGCAGACAGTGGCGGGGATATCCGGCTCCCAATCCGGAATCT  1324
Qy      307  -----Ser-P  308
```

```
Db 1325 CTCTCTCCGGTTACGAAGAGTGTGAGCAAGAAATTCGACAGACTCAGCTCAGCTCAGC 1384
QY 308 rGlnProGln-----AlaValAlaValProA 317
Db 1385 CTCAGCCACAGATTCAAGCTCAAGTTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1444
QY 317 lAlys-----ArgLysAspLeuHisMetLeuValTTPSerS 329
Db 1445 CCAACAGTGGGAACAAGCCACCAGCTGTAGAGAGTTCACATGTTGTGTGGAGCT 1504
QY 329 eSerAlaSerProValSerGluArgAlaAlaValHisValPhe---GlyAla-----G 346
Db 1505 CCAGCAGCTGCGCGGTCTCCGAAGCGCGGCTCCAGCTGTTAGTGGGCTGATTTCG 1564
QY 346 lGlyAlaAspHisAlaAspValLeuAlaLysGlyValAlaGln----- 359
Db 1565 GAGCGTCTCAATCCGCTCGCTCAGAACAGGGGCCAAGAGATTAGGATTGGTGTG 1624
QY 360 -----AlaTyrAspGluTyrGlyA 366
Db 1625 CTGATGACCACCTCAAAATGGAAGAAACCAACAAAGCTGCAGCAGAGAGATTGGTG 1684
QY 366 rGAspAspTyrSerSerArgThrLysAsnGlySerGlyGly-----AlaAspLysG 383
Db 1685 GTGAAGATTGAATTTCCAGTTAAGAGGAGAGAACAGCAGATGAGAGGAGGAGAAAG 1744
QY 383 lGlyPro---ThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysA 402
Db 1745 CAGGACCGGTGGTCTTAACAGAGCTTGGATCTAGTTCACGCGGAGGCTACACCGAAT 1804
QY 402 sAspGlyGluGlyArgAlaAlaAlaValAla-----MetProAlaSerValM 419
Db 1805 CC-----GCCGTGCGGTGGCTGGAAGAACACATGCTCCGCGGAGTGTA 1849
QY 419 eThrArgLeuLeuLeuMetValTTPArgLysLeuLeuLeuArgAsnProAsnThrTyrS 439
Db 1850 TGACTCGTCTATCTGATTATGTTGGTGGAGAGAGCTTATCCGCAACCCCAACACTTACT 1909
QY 439 eSerLeuLeuGlyValValTTPSerLeuValSerTyrArgTTPGlyLeuMetProA 459
Db 1910 CTAGCCTATTGGTGTCTTGGTCCCTCGTGGCGTTTAGGTGGCATGTGCAAAATGCCCA 1969
QY 459 lAlleAlaArgSerLysLeuSerAspAlaGlyLeuMetAlaMetPheS 479
Db 1970 AAATATAGAGAAATCAATCTCCACTGCTGATGCTGGTCTTGGAAATGCTATGTTCA 2029
QY 479 eLeuGlyLeuPheMetAlaLeuGlnProArgLysLeuAlaCysGlyAsnLysLeuAlaA 499
Db 2030 GCTTGGGTCTGTTTCATGGCTCTTCAACCTAAGATAATTGCTAGTGGGAACCTGTGGCA 2089
QY 499 lAlleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerLysA 519
Db 2090 CATTGGCCATGCTGTGATTCCTCAACAGGTCGCGGCTCATGGCAGCAGCTTCCATCG 2149
QY 519 lValGlyLeuArgGlyValLeuLeuHisLysLeuAlaValGlnAlaAlaLeuProGlnG 539
Db 2150 CTGTTGGCTACGTGGCACCCCTTACGTGTAGCTATTGTTCCAGCTCAGCTTCCATCG 2209
QY 539 lYleValProPheValPheAlaLysGlyTyrGlyValHisProAspLysLeuSerThra 559
Db 2210 GGATTGTTCCATTGTTGTTGTTAGGAGTAAACGCTCCACCCAGCCATTCTTAGTACAG 2269
QY 559 la-----TyrGly-----ProIleThr 564
Db 2270 CGGTTATATTGGGATGTTGATAGCACTACCAATTACT 2307
```

## RESULT 5

```
US-10-260-238-30
; Sequence 30, Application US/10260238
; Publication No. US20040016025A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, Paul R.
; APPLICANT: Moughamer, Todd G.
```

```
; APPLICANT: Briggs, Steven P.
; APPLICANT: Cooper, Bret
; APPLICANT: Glazebrook, Jane
; APPLICANT: Goff, Stephen A.
; APPLICANT: Ktagiri, Fumiaki
; APPLICANT: Kreps, Joel
; APPLICANT: Provart, Nicholas
; APPLICANT: Ricke, Darrell
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION
; FILE REFERENCE: 60111-NP
; CURRENT APPLICATION NUMBER: US/10/260,238
; CURRENT FILING DATE: 2002-09-26
; PRIOR APPLICATION NUMBER: US 60/325,448
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/325,277
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/370,620
; PRIOR FILING DATE: 2002-04-04
; NUMBER OF SEQ ID NOS: 6077
; SEQ ID NO 30
; LENGTH: 1759
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1584)..(1584)
; OTHER INFORMATION: n = any nucleotide
; FEATURE:
; NAME/KEY: N region
; LOCATION: (1587)..(1587)
; OTHER INFORMATION: n = any nucleotide
; NAME/KEY: N region
; LOCATION: (1602)..(1602)
; OTHER INFORMATION: n = any nucleotide
; US-10-260-238-30
```

```
Alignment Scores:
Pred. No.: 7,53e-157 Length: 1759
Score: 1603.50 Matches: 373
Percent Similarity: 68.74% Conservative: 47
Best Local Similarity: 61.05% Mismatches: 96
Query Match: 55.10% Indels: 96
DB: 15 Gaps: 21
```

US-10-030-884-14 (1-573) x US-10-260-238-30 (1-1759)

```
QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20
Db 1 ATGATAACGGGGCGGCGACTTCTACACAGTGATGACGGCGATGGTGGCGTACGTGGCG 60
QY 21 MetThrLeuAlaTyrGlySerValArgTTPArgLysPheThrProAspGlnCysSer 40
Db 61 ATGATCCTGGCGTACGGGTGGTGAAGTGGTGGCGATCTTCACGCCCGCACCATGCTCC 120
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
Db 121 GGGATCAACCGCTTCGGGGCTCTTCGGCGTCCCGCTCTCTCTCTCTCTCTCTCTCTCC 180
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
Db 181 ACCAACAAACCCCTACACCATGAACCTCCGCTTTCATCGCGCGCGACACCCCTGCAGAAAGCTC 240
QY 81 AlaValLeuAlaLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100
Db 241 ATGCTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 291
QY 101 LeuAspTTPSerIleThrLeuPheSerLeuSerThrLeuProAsnThrLeuValMetGly 120
Db 292 CTCGAGTGGACCATCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 351
QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140
```





Percent Similarity: 63.87% Conservative: 73  
Best Local Similarity: 52.92% Mismatches: 102  
Query Match: 53.63% Indels: 139  
DB: 9 Gaps: 20

US-10-030-884-14 (1-573) x US-09-938-842A-847 (1-1923)

QY 1 MetIleThrAlaLeuAAspLeuThrHisValLeuThrAlaValProLeuThrValAla 20  
DB 1 ATGATCTCATGCGACGACTCTACACGGTCTCTCAACCGCGGTGATTCCTCTCTACGTGGCC 60

QY 21 MetThrLeuAlaTyGlySerValArgTrpArgIlePheThrProAspGlnCysSer 40  
DB 61 ATGATCTCTCGTTACGGTCTCTCGGTGGTGGAAATCTCTCACCGACCAATGTCTCC 120

QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
DB 121 GGAATCAACCGTTTGTGGCATCTCTCGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180

QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
DB 181 ACCAACAATCTTCACGCAATGATCTCCGGTTCATCGCGCGGACACTCTCAAAATATC 240

QY 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100  
DB 241 ATCATGTTGTCACTTTTATGCTCTATGGCT-----AATTCACCTCGCTCCGGTAGC 291

QY 101 LeuAspTrpSerIleThrLeuPheSerLeuSerThrIleuProAsnThrLeuValMetGly 120  
DB 292 CTCGAGTGAGCATCACAACTCTTTTCTCTCTCCACACTCCACACTCTCTCTCTCTCTCT 351

QY 121 IleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140  
DB 352 ATTCCTCTCTGATCGCATGATGGCAATCTCT---GGTTCCTCATGTGTCACAAATC 408

QY 141 ValValLeuGlnCysIleIleTrpThrLeuMetLeuPheLeuPheGluTrpArgAla 160  
DB 409 GTCTGCTCTCAGTGATCATCTGGTACACGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 468

QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180  
DB 469 GCGAGATCTCTATCATGAGCAGTCTCCCTGAG---ACGGCTGCTTCCATGTTCTTCTTC 525

QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAsp---ValGluLeuGluAlaGln 199  
DB 526 AAAGTCGAATCGAGCTGCTTTCTGTCGACGCCATGATTTCTTGACCGCATGACGAG 585

QY 200 ProAspGlyValAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal 219  
DB 586 ATA-----GGTGACGACGGGAGCTT 606

QY 220 ArgValThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSerHis 239  
DB 607 CACGTCAACCGTGAGAAATCCACAGCTTCACTG---CGTTCGTTCTGC----- 651

QY 240 SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleThrSerLeuGln 259  
DB 652 GGCCTGAAATGACTCCACGCGCGTCAAAATCTCACCGGAGCTGAGATTATAGTCTCAGC 711

QY 260 SerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPhePheIle 279  
DB 712 ACC-----ACTCTAGAGGCTCTAATTCACCACTCTGATTTTATACACATG 759

QY 280 ValGly-----Ala 282

DB 760 ATGGGTTTCCCGGTGGTGTCTCTCCAATTTGGTCCGGCGGATATGACTCCGTTCAA 819

QY 283 AlaAlaLysGly-----GlyGlyGlyAlaAlaGly 292

DB 820 TCATCTAGAGTCCAACTCTCGACCTTCAAACCTTCGAGAGAAATTCGCCATGCCATCC 879

QY 293 AspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 310

DB 880 TCCCCGAGATTCCGGTATTACCTCGGAGGAGGCGGGTCTTATCCGGCTCCGATCCA 939

QY 310 ----- 310

DB 940 GAGTTCTCTTCAACCAACCATCTACCGCAATAAAGCGTCAATAAAAAACCGAAGAC 999

QY 311 -----GlnAlaValAlaValProAla-----LysArg 319

DB 1000 GTTAATACGATCAGCAGAGGACTCTTCAACGGGCGGCAAGTCAACACGATGACGCC 1059

QY 320 LysAspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAla 339

DB 1060 AAGGAGCTTCAATGTTCTGTGGAGCTCAACCGGTCTACCGGTTTTCAGACCGGGGGT 1119

QY 340 ValHisValPheGlyAlaGlyGlyAlaAspHis-----AlaAspValLeu 354

DB 1120 CTTAAGCTTTTTCGC---GGAGCACTGCAACGATCAAGCGGAGCACTGACCAAGT 1176

QY 355 AlaLys----- 356

DB 1177 GCTAAAGAGATCGTATGTTAGTCCAGATCAATCTCACAAACGGCGAGCAACGCTGTA 1236

QY 357 -----GlyAlaGlnAlaTyAspGluTyArgValArgAsp 367

DB 1237 GCTCATCCAGCAAGTGGAGATTTCCGAGGAGAAACAATTTAGTTTCCCGGAAAGAA 1296

QY 368 AspTySerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeu 387

DB 1297 GAGAAGCAGAGAGACCAAAAGACCGCGAGATGTT-----CTA 1335

QY 388 SerLysLeuGlySerAsnSerThrAlaGlnLeuTyProLysAspAspGlyGluGlyArg 407

DB 1336 AACAACTTCTCTCCAAATTTCCACGGCGGCTCAATCAAG---ACAGCTTAGAGGA 1392

QY 408 AlaAlaAlaValAla-----MetProAlaSerValMetThrArgLeuLeuLeu 424

DB 1393 GCGAGCAAGTCAACGAAAAATATGCTCCCGCGAGTGTGATGACAAAGCTGATATC 1452

QY 425 IleMetValTrpArgLysLeuIleArgAsnProAsnThrTySerSerLeuIleGlyVal 444

DB 1453 ATAATGTTTGGAGAAACTCATCAGAAACCCAAACACTTACTCTACTCTCATTTGACCT 1512

QY 445 ValTrpSerLeuValSerTyArgTrpGlyIleGluMetProAlaIleIleAlaArgSer 464

DB 1513 ATTTGGCTCTCTCGCTTTCGGTGGCAGTGGCAATGCCAAATCATTCACCAATCT 1572

QY 465 IleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMet 484

DB 1573 ATCTCCATCTATCTGATGCTGCTCTTGGAAATGCAATGTTTGTGGGTTTGTTCATG 1632

QY 485 AlaLeuGlnProArgIleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyVal 504

DB 1633 GCGTTGCAACCCAAATTAATCGTTTGGGAATTCAGTGGCAACGTTTGCATGGCGGTT 1692

QY 505 ArgPheValAlaGlyProAlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGly 524

DB 1693 AGGTTCTTACGGTCCGGCGGTGATGGCGGTGCTGTATAGCCATCGGATTCAGTGGT 1752

QY 525 ValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheVal 544

DB 1753 GATTACTCGGTGTCGTATAGTTTACGCGGCAATCTACCTCAGGAATTTGCCCTTTGTG 1812

QY 545 PheAlaLysGluTyGlyValHisProAspIleLeuSerThrAla-----TyGly--- 561

DB 1813 TTTCCGAGGAGTACATGTTTCATCTCTCTCTTAAAGTACAGGGTAAATTTTGAATG 1872

QY 562 -----ProIleThr 564

DB 1873 CTTATAGCGTTCCGATCAGC 1893

RESULT 8  
US-09-938-842A-847  
; Sequence 847, Application US/09938842A

QY	200	ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal	219
DB	586	ATA-----GGTGAACGACGGAGCTT	606
QY	220	ArgValThrValArgLysSerThrSerArgSerGluAlaAlaCysSerHisSerHis	239
DB	607	CACGTACCGTGAGAAAAATCAACGGCTCACGT---CGTTCTGTTCTGC---	651
QY	240	SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGln	259
DB	652	GGCCGGAACATGACTCCACGGCGCTCAAAATCTCACCGAGCTGAGATTTATAGTCTCAGC	711
QY	260	SerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaPhePheAsnIle	279
DB	712	ACC-----ACTCTAGAGGCTCTAAATTTCAACCACTCTGATTTTTTACAACATG	759
QY	280	ValGly-----Ala	282
DB	760	ATCGGTTCCCCGGTGGTGTCTCTCAATTCGGTCCGGCGGAGATGTACTCGGTTCAA	811
QY	283	AlaAlaLysGly-----GlyGlyGlyAlaAlaGly	292
DB	820	TCATCTAGAGGTCCAACTCTCTCGACCTTCAAACTTCGAGGAGAATTCGCCATCGCATCC	879
QY	293	AspGluGlnLysGlyAlaCysGlyGlyGlyGlyGlyHisSerProGlnPro-----	310
DB	880	TCGCCGAGATTCGGGTATTACCTCGGAGGAGACCGGGTCTTATCCGGCTCCGAATCCA	939
QY	310	-----	310
DB	940	GAGTTCCTTCACCAACCACCATCTACCGCCAAATAAAGCGTCAATAAAACCCGAAAGAC	999
QY	311	-----GlnAlaValAlaValProAla-----LysArg	319
DB	1000	GTTAATACGAATCAGCAGCAGCTCTCTTCAACGGCGCGCAAGTCAAAACAGCCATCGAGGCC	1059
QY	320	LysAspLeuHisMetLeuValTrpSerSerAlaSerProValSerGluArgAlaAla	339
DB	1060	AAGGAGTTTCACATTTCTGCTGAGCTCAACGGGTACCCGTTTCAGACCGCGGGGT	1119
QY	340	ValHisValPheGlyAlaGlyGlyAlaAspHis-----AlaAspValLeu	354
DB	1120	CTTAAACGTTTTCGSC---GGAGCACCTTGACAAACGATCAAGCGCGAAGATCTTGACCAAGGT	1176
QY	355	AlaLys-----	356
DB	1177	GCTAAAGAGATCCGTATGTATTGCCAGATCAATCTCAACAGCGGAGACCAAAAGCTGTA	1236
QY	357	-----GlyAlaGlnAlaTyrAspGluTyrGlyArgAsp	367
DB	1237	GCTCATCCAGCAAGTGGAGATTTCGGAGGAGAACACAAATTTAGTTTCCCGCGAAAGAA	1299
QY	368	AspTyrSerSerArgThrLysAsnGlySerGlyGlyAlaAspLysGlyLysProThrLeu	387
DB	1297	GAAGAAGCAGAGAGACCAAAAGACCGCGAGAATCGT-----CTA	1339
QY	388	SerLysLeuGlySerAsnSerThrAlaGlnLeuTyrProLysAspAspGlyGluGlyArg	407
DB	1336	AACAAACTTGCTCCAAATTCACCGCGCGCGCTACAATCCAG---ACAGGCTCAGGAGGA	1399
QY	408	AlaAlaAlaValAla-----MetProProLysValMetThrArgLeuIleLeu	424
DB	1393	GCCGAAGCAAGTCAACGAAAAAAATATGCCTCCGCGAGTGTGATGACAGGCTGATACGTG	1459
QY	425	IleMetValTrpArgLysLeuIleArgAsnProLeuThrTyrSerSerLeuIleGlyVal	444
DB	1453	ATAATGTTGGAGGAACATCATCAGAAACCCAAACACTTACTCTAGTCTCATTTGGACTT	1511
QY	445	ValTrpSerLeuValSerTyrArgTyrGlyIleGluMetProAlaIleIleAlaArgSer	464
DB	1513	ATTTGGGCTCTCGTCGCTTTCGGTGGCAGCTGGCAATGCCAAAAATCATTCAACAATCT	1577

```
QY 465 ILeSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMet 484
Db 1573 ATCTCCAACTTATCTGATGCTGGTCTTGGAAATGCAATGTTAGTTTGGGGTGTTCATG 1632
QY 485 AlaLeuGlnProArgIleIleAlaCysGlyAsnLysLeuAlaIleAlaMetGlyVal 504
Db 1633 GCGTTGCAACCAAAATTAATCGCTTGTGGGAATTCAGTGGCAACGTTTGGCATGGCGGTT 1692
QY 505 ArgPheValAlaGlyProAlaValMetAlaAlaSerIleAlaValGlyLeuArgGly 524
Db 1693 AGGTTCCCTTACGGTTCGGCGGTGATGCGGTTGCTGTATAGCCATCGGAATTACGTGGT 1752
QY 525 ValLeuLeuHisIleAlaIleValGlnAlaAlaLeuProGlnGlyIleValProPheVal 544
Db 1753 GATTTACTGCGTGCCTATAGTTTCAAGCGCAATTAATCTCAAGAAATGTGTGCCCTTTGTG 1812
QY 545 PheAlaLysGlyIleGlyValHisProAspIleLeuSerThrAla-----TyrGly--- 561
Db 1813 TTTCGAAGGAGTACAATGTTCATCCTGCTATTTTAAGTACAGGGTAAATTTTGAATG 1872
QY 562 -----ProIleThr 564
Db 1873 CTTATAGCGCTTCGATCAAG 1993
RESULT 9
US-09-887-576-628
; Sequence 628, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 628
; LENGTH: 2222
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-628
Alignment Scores:
Pred. No.: 2,386-150 Length: 2222
Score: 1542.50 Matches: 349
Percent Similarity: 62.58% Conservative: 59
Best Local Similarity: 53.53% Mismatches: 111
Query Match: 53.01% Indels: 133
DB: 9 Gaps: 17
US-10-030-884-14 (1-573) x US-09-887-576-628 (1-2222)
QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
Db 1 ATGATTACGGTGGCAGACTTTGACCGTCTTCACCGCGGTGGTACCACTTTACGTAGCT 60
QY 21 MetThrLeuAlaTyrGlySerValArgTrpTrpArgIlePheThrProAspGlnCysSer 40
Db 61 ATGATTCTCGCTACGGATCCGTACAGTGGTGAAGATATTTCTCACCAGACAGATGCTCC 120
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
```

```
Db 121 GGCATCAACCGCTTCGTGCTATATTCGCGGTCCTCTCTCTCCATCTCATCTCC 180
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
Db 181 ACCAACGATCCTTACGCCATGAATTCGCTTCGCGCCGACACGCTTCCAAAATC 240
QY 81 AlaValLeuAlaLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100
Db 241 ATCATGCTCGTCTTACTTCTCTATGGGCT-----AACCTAACCAAGAACGGTAGC 291
QY 101 LeuAspTrpSerIleThrLeuPheSerThrLeuProAsnThrLeuValMetGly 120
Db 292 TTGGAGTGATGATCACAACTCTCTCTCAGCATCTCCCAACACACTCTTGTCTATGGG 351
QY 121 IleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeuMetValGlnVal 140
Db 352 ATCCCTCTGTTGATGCCATGTACGGA---ACCTACGCAAGGTCTCTAATGGTCCAAAGTC 408
QY 141 ValValLeuGlnCysIleIleTrpTyrThrLeuMetLeuPheLeuPheLeuTyrArgAla 160
Db 409 GTTGTCTTTCAGTGATCATTTGGTACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 468
QY 161 AlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSerIleValSerPhe 180
Db 469 GCTAAGCTTCTTATTATGGAGCAGTTCGCGAG---ACTGGTGCCTTATTGTTTCGTTT 525
QY 181 ArgValAspSerAspValValSerLeuAlaArgGlyAsp---ValGlnLeuGlnAlaGlu 199
Db 526 AAAGTTGAATCCGACGCTGTTTCTCTCGACGCTCATGATTTTCTTGAGACGATGCTGAG 585
QY 200 ProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAspAlaGlyArgVal 219
Db 586 ATA-----GGAAACGACGGGAAGCTT 606
QY 220 ArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCysSerHisSerHis 239
Db 607 CATGTTACCGTGAGGAATCAAAACGATCGAGA-----CGG 642
QY 240 SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGln 259
Db 643 TCACTGATGACTCTCCAGCGCTTCAATCTTACCGAGCTGAGATCTATAGTCTTAGT 702
QY 260 SerSerArgAsnProThrProArgLysSerSerPheAsnHisAlaAspPhePheAsnIle 279
Db 703 TCG-----ACTCCGAGAGGTTCTAACTTTAACCAITCTGATTTTACTCTGTT 750
QY 280 ValGlyAlaAlaLysGlyGlyGlyAlaAlaGlyAsp----- 293
Db 751 ATGGGGTTTCCC-----GGCGGAGGCTTTTCGAATTTTGTGTCGCGGATTTG 798
QY 294 -----GluGlyLysGly 297
Db 799 TACTCCGTTCAATCTTCTCGTGGTCCGACTCCACGCGCTTCGAACTTCGAAGAGAACAC 858
QY 298 AlaCysGlyGlyGly-----GlyGlyGlyHisSerProGlnProGlnAlaValAlaVal 315
Db 859 GCCGTTAAATATGGAATTTACAATAACACTAACAAGATTCTGTTCCGCGCGCGGTTCTG 918
QY 316 ProAlaLysArg----- 319
Db 919 CCGGCTCCGACCCCGAGTTTTCACCGGTACGGGTGTTTCAACTAAACCGAATAAAATT 978
QY 320 -----Lys 320
Db 979 CCTAAAGAAACCAACAGCAACTGCAAGAGAAAGATAGCAAGCGAGCCATGACGTAAG 1038
QY 321 AspLeuHisMetLeuValTrpSerSerSerAlaSerProValSerGluArgAlaAlaVal 340
Db 1039 GAGCTTCACATGTTTGTGGAGCTCAACGGCTTCTCCGGTCTCCGAC----- 1086
QY 341 HisValPheGlyAlaGlyAlaAspHisAlaAspValLeuAlaLysGlyAlaGlnAla 360
```

```
Db 1087 ---GTGTTTGGCGAGGTGACGGACACACGTG-----GCAACGGAA 1125
Qy 361 TyrAspGluTyrGlyArgAspTyr-----SerSerArgThrIys 374
Db 1126 CAATCTGAACAAGGTGCTAAGAGATTGGATGGTGTCTCTGATCAACCTCGAAGAGT 1185
Qy 375 AsnGlySerGlyGlyAlaAsp----- 381
Db 1186 AATGCTAGAGGTGGTGTGATGATATCGCGGTCTTGATAGTGAGAGGAGAAAGAG 1245
Qy 382 -----LysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAlaGlnLeuTyr 399
Db 1246 ATAGAGAAAGCTACAGAGGCTGAATAAATGGGGTCTTAATTCACCGCGGAGTAGAG 1305
Qy 400 ProLysAspAspGlyGlyAlaAlaValAlaMetProAlaSerValMet 419
Db 1306 GCGGCTGGTGGAGATGCGCGGCGCAACACGACACATATGCGCGCGACAAAGTGTGATG 1365
Qy 420 ThrArgLeuLeuLeuMetValTyrArgLysLeuLeuLeuAsnProAsnThrTyrSer 439
Db 1366 ACACGACTGATATTGATAATGTGTGGAGAAAGCTGATCAGAAACCCAAACACGTACTCC 1425
Qy 440 SerLeuLeuGlyValValTyrSerLeuValSerTyrArgTyrGlyLysLeuMetProAla 459
Db 1426 AGTCTAATCGGCTCATATGGGCTCTTGTGCTTACCGGTGCGCATGTGCTATGCCAAA 1485
Qy 460 IleLeuAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAlaMetPheSer 479
Db 1486 ATATTACAAATCCATCTCCATCTCTCAGATGCTGTGCTTTGGAATGGCTATGTTACG 1545
Qy 480 LeuGlyLeuPheMetAlaLeuGlnProArgIleLeuAlaCysGlyAsnLysLeuAlaAla 499
Db 1546 TTAGGTTTATCATGGCACTTCAACCAAAATCATTGCTTGTGGNATCTGTGCCACG 1605
Qy 500 IleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaSerIleAla 519
Db 1606 TTGGCATGGCGTCAGATTATATACCGGTCCGGGCATCATGGCTGTGGATTGCC 1665
Qy 520 ValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGlnAlaLeuProGlnGly 539
Db 1666 ATGGGTTACACGGCGACCTTCTCGTATAGCATCTGTTACGGTGGCTGCTGCTCAAGA 1725
Qy 540 IleValProPheValPheAlaLysGluTyrGlyValHisProAspIleLeuSerThrAla 559
Db 1726 ATAGTTCCGTTGTGTTGCAAAAGAGTACAATGTGCATCCACGATCTAAGCACTGGG 1785
Qy 560 -----TyrGly-----ProIleThr 564
Db 1786 GTCATATTGGAAATGTTAATAGCCTTACCTATAACT 1821
```

## RESULT 10

```
US-10-425-114-1349
; Sequence 1349, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 1349
; LENGTH: 2876
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700151120_FLI
```

```
US-10-425-114-1349
Alignment Scores:
Pred. No.: 2,78e-150 Length: 2276
Score: 1542.00 Matches: 357
Percent Similarity: 65.61% Conservative: 57
Best Local Similarity: 56.58% Mismatches: 135
Query Match: 52.99% Indels: 82
DB: 12 Gaps: 19

US-10-030-884-14 (1-573) x US-10-425-114-1349 (1-2276)
Qy 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
Db 115 ATGATCACCAGCGCGGACATCTACAGCTTCTTGGCGCATCGTGCCTGTACGTGGCC 174
Qy 21 MetThrLeuAlaTyrGlySerValArgTyrPheArgIlePheThrProAspGlnCysSer 40
Db 175 ATGTTCTGGCGTACGGTCCGTGCGGTGGTGGCCATCTTACCCCGGACCACTGCTCC 234
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
Db 235 GGCATCAACCGCTTCTGGCGGCTTCTTGGCGGCTGCTTCTTCTTCCACTTCATCTCC 294
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
Db 295 AGCTGGACCCCTACGCGATGACGAGTACCGGTTCTTGGCGCGCATCTCGTGCAGAAAGCTG 354
Qy 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArg 97
Db 355 GTCATCTGGCGCGCTGGCGGCTGGCACACGCTCTCTCCGCTACCGCGGCGGCGG 414
Qy 98 ---AlaLeuGlyLeuAspTyrPheIleThrLeuPheSerLeuSerThrLeuProAsnThr 116
Db 415 GCGGCGTCTGCTGGTACGACCATCACGCTCTTCTGCTGTCCACGCTGCTCCCAACAG 474
Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerIleGlyThrLeu 136
Db 475 CTGCTGATGGCATCCCGCTGCTCCGCGCATATACGGGACTTCTCG---GGCAACCTC 531
Qy 137 MetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156
Db 532 ATGTGTCAGATCTGTGTGCTGCGAGCGTCATCTGTACAGCTCATGCTCTTCTTCTTCT 591
Qy 157 GluTyrArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaSer 176
Db 592 GAGTACCGCGCGCAAGGCGCTCATCTCGAGCAGTTCCCGCGCGACGCTCGCGCGCAGC 651
Qy 177 IleValSerPheArgValAspSerAspValValSerLeuAlaArgGlyAspValGluLeu 196
Db 652 ATCGCTCTCTTCAAGGTCGACTCCGAGCTGCTCTCTCTC---AACGGCGGAGGCGCTG 708
Qy 197 GluAlaGluProAspGlyValAlaGlyAlaValSerSerArgGlyGlyAspAla 216
Db 709 CAGCGCGAC-----GCCGAGGTCCGCGACGCGAC 735
Qy 217 GlyArgValArgValThrValArgLysSerThrSer----- 228
Db 736 GCGCGCGTCCACGCTCATCCCGCGCTCCGCTCCGCTCCACCGCGCGCGCGCGCAC 795
Qy 229 -----SerArgSerGluAlaAlaCys---SerHisSerHisSerGlnThrMetGlnPro 245
Db 796 GCGCGCGCGCTCCGCGCGCGGTACCGCCCTACGCGCGCTCTCGGCGCATGACCGCG 855
Qy 246 ArgValSerAsnLeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThr 265
Db 856 CGCGCTCCCAACCTCACCGCGCTGAGATCTACTGCTGAGACCTCGCGGAGCCCGACG 915
Qy 266 ProArgGlySerSerPheAsnHisAlaAspPheAsnIleValGlyAla----- 282
Db 916 CCCCGCGCTCCAGCTTCAACGACGCTTCTTACGCCATGTTCAACGGGAGCAAGATG 975
Qy 283 -----AlaAlaLysGlyGlyGlyAlaAlaGly---AspGluGlu----- 295
```



```

586 ATC-----GGTGACGATGGTAAGCTT 606
220 ArgValThrValArgLysSerThrSerArgSerGluAlaAalaCysSerHisSerHis 239
607 CARGTTACGGTGAGAAATCAACGGCTTCTCGG-----AGATCTCTTTTACGGTGGTGT 660
240 SerGlnThrMetGlnProArgValSerAsnLeuSerGlyValGluLeuIleTyrSerLeuGln 259
661 GGTAATAATAGTACTCTCGTCGCTCTAATCTACCGGAGCTGAGATTATAGTCTT-- 717
260 SerSerArgAsnProThrProArgGlySerSerPheAsnHisAlaAspPheAsnIle 279
718 -----AACACTACTCCAGAGAGTCTAACTTCAATCATCTGATTTTACTCTATG 768
280 ValGly-----Ala 282
769 ATGGGGTTTCCGGTGGCGGGCTTTTCGAACCTTGGTCCGGCGGATATGACTCTGTTGAG 828
283 AlaAlaLysGly-----GlyGlyAlaAlaGly 292
829 TCCTCGAGAGGACTTACTCTAGACCTTCGAATTTTGGAGAGAGTTGTGCTATGCTTCT 888
293 AspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyHisSerProGlnProGln-- 311
889 TCGCCAAGATTCGGGTATTACCGGAGGAGCTCCCGGTCGTACCCAGCTCTTAATCCG 948
312 -----AlaValAlaValProAlaLysArg----- 319
949 GAGTTTTTCTACCGGTAATAAAACCGGTAGTAAGCTCCAAAGAGAAATCATCATGTGTA 1008
320 -----LysAspLeuHisMetLeuValTrpSerSerAla 331
1009 GGAATATCAATAGTATGATGCTAAGGAGCTTCATATGTTTGTGGGATCCACCGA 1068
332 SerProValSerGluArgAlaAlaValHisVal-----PheGlyAla----- 345
1069 TCACCGGTTTCGACCGAGCTGCTCTTCAAGTTGATAATGAGGCAATCAACAAGTCGA 1128
346 -----GlyGlyAlaAspHisAlaAspValLeuAla----- 355
1129 AAATCCGATCAAGCGCGTCCAAAGAGATTCGAATGCTGATCTGATCATACTCAAAAT 1188
356 -----LysGlyAlaGlnAlaTyrAspGluTyrGlyArgAspAspTyrSerSerArg 372
1189 GGTGAAACAAAGCTGGTCCGATGAACGGGACTATGGCGGGAAGAGAGTCGGAGAGG 1248
373 ThrLysAsnGlySerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySer 392
1249 GTAAAGGAGAGTGCCTAACGGA-----CTACACAAGCTTCGGTGT 1287
393 AsnSerThrAlaGlnLeuTyrProLysAspAspGlyGlu-----GlyArgAlaAlaVal 411
1288 AATCCACAGCAGAGCTTAACCTTAAGAGCTATAGAAACGGTGAAACTGTACCGGTA 1347
412 Ala---MetProAlaSerValMetThrArgLeuIleLeuIleMetValTrpArgLys 430
1348 AAACATATGCCACAGCGAGTGATGACTCGGCTGATATTGATATGTTGGTGGAGAAA 1407
431 LeuIleArgAsnProAsnThrTyrSerSerLeuIleGlyValValTrpSerLeuValSer 450
1408 CTATAGAAACCAACACTTACTCTAGTCTCATTTGGTCTCATTTGGGCTCTTGTGCT 1467
451 TyrArgTrpGlyIleGluMetProAlaIleAlaArgSerIleSerIleLeuSerAsp 470
1468 TTCAGTGGGATGTGGCAATGCTCAAAATTAATCAACAATCAATCTCAATCTTCTGAT 1527
471 AlaGlyLeuGlyMetAlaMetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIle 490
1528 GCTGGTCTCGTATGGCAATGTTTCAGTTGGGGTGTGTTATGCGATTCGACCGCAATTA 1587
491 IleAlaCysGlyAsnLysLeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyPro 510
1588 ATTGCTTGGGAATTCGACGGCGACTTTTGGGATGGCGGTGAGATTCTTTACTGACCA 1647

QY 511 AlaValMetAlaAlaAlaSerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAla 530
Db 1648 GCGTAATAGCGCGTGGCAGCAATGGCTATTGGATTACGTGAGACCTATTGCGTGGCC 1707
QY 531 IleValGlnAlaAlaLeuProGlnGlyIleValProPheValPheAlaLysGluTyrGly 550
Db 1708 ATTGTTCAAGTCGATTCCTCAAGGATCGTCCGTTTGTGTTTGCAAAAGAGTATAAT 1767
QY 551 ValHisProAspIleLeuSerThrAla-----TyrGly-----ProIle 563
Db 1768 GTTCATCCCGCAATCTTGAGTACAGGGTAAATTTTGAATGCTTATCGCACTACCGATT 1827
QY 564 Thr 564
Db 1828 ACA 1830

RESULT 12
US-09-938-842A-848
; Sequence 848, Application US/09938842A
; Publication No. US20040009476A9
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 848
; LENGTH: 1860
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-848

Alignment Scores:
Pred. No.: 9,64e-149 Length: 1860
Score: 1536.00 Matches: 346
Percent Similarity: 64.59% Conservative: 68
Best Local Similarity: 53.98% Mismatches: 119
Query Match: 52.44% Indels: 108
DB: 11 Gaps: 19

US-10-030-884-14 (1-573) x US-09-938-842A-848 (1-1860)
QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValProLeuTyrValAla 20
Db 1 ATGATCATATGGCAGCACCTCTACCGTCTCTACGGCGGTGATACCACTCTACGTAGCC 60
QY 21 MetThrLeuAlaTyrGlySerValArgTyrTrpArgIlePheThrProAspGlnCysSer 40
Db 61 ATGATCTTCGTACCGCTCAGTCCGGTGGTGGAAATCTTCTCACCAGACCAATGTTC 120
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60
Db 121 GGCATAAACCGTTCGTCGCTATTTTCGCGTCCCTCTCTCTCTTCTCTCTCTCTCTCC 180
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80
Db 181 TCAAACCAATCTTACCGCAATGATCTCCGATTCTATCGACCGCATACCTCCAAACTA 240
QY 81 AlaValLeuAlaLeuLeuAlaLeuAlaSerArgGlyLeuSerSerProArgAlaLeuGly 100
Db 241 ATCATGCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 291
```



DB: 9 Gaps: 15  
US-10-030-884-14 (1-573) x US-09-938-842A-1305 (1-1983)  
QY 1 MetIleThrAlaLeuAspLeuTyrHisValLeuThrAlaValValProLeuTyrValAla 20  
DB 1 ATGATCACCGGCAAGACATGATACGATGTTTAGCGGTATGTCGCGCTATACGTTGCT 60  
QY 21 MetThrLeuAlaTyrGlySerValArgTyrTpAlaGlyPheThrProAspGlnCysSer 40  
DB 61 ATGATATTAGCTATGTTTCGTTACGTTGTTGGGGATATTCACACCGGACCAATGTTCC 120  
QY 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
DB 121 GGTATAACCGGTTCTGCGGTTTTCGCGGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
QY 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
DB 181 TCCATGATCCCTATGCAATGATATACCACTTCTCTGCTGCTGATCTCTTCAGAAATGC 240  
QY 81 AlaValLeuAlaLeuLeu-----AlaLeuAlaSerArgGlyLeuSerSerPro 96  
DB 241 GTTATCTCGCGCACTCTTCTTTTGGCAGGCGTTTAGCGCGAGAGGA-----288  
QY 97 ArgAlaLeuGlyLeuAspTyrPheIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
DB 289 -----AGCCTAGATGGATGATACGCTCTTTTCACTATCAACACTGCTCAACACG 339  
QY 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyrGlyAlaSerSerAlaGlyThrLeu 136  
DB 340 TTGGTAATGGGAATCCATTGCTTAGCGGATGTACGGAGACTTCTCC---GGTAACCTA 396  
QY 137 MetValGlnValValLeuGlnCysIleIleTyrThrLeuMetLeuPheLeuPhe 156  
DB 397 ATGGTTCAGATCGTGTGTTCCAGACATCATGATGATATACATTAATGCTCTCTCTGTT 456  
QY 157 GluTyrArgAlaAlaArgAlaLeuLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
DB 457 GAGTTCGTTGGGTAAAGTCTTCATCTCCGAGCAGTTCCTCGAG---ACGCTGGTTCA 513  
QY 177 IleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGlu 195  
DB 514 ATTACTTCTCTCAGAGTTGACTGTGATGTTATCTCTCTTAATGGCGGTGAACCCCTCCAG 573  
QY 196 LeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAsp 215  
DB 574 ACCGATCGCGAGATA-----GGAGAC 594  
QY 216 AlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys 235  
DB 595 GACGGAAGCTACACGCTGTGTTGAGAGATCAAGTGGCGCTCATCATGATGATCTTTCA 654  
QY 236 -----SerHisSer-----HisSerGlnThrMetGlnProArgValSerAsn 249  
DB 655 TTCACAAATCTCACGGCGAGACTTAATCTCTCCATGATAAACCGCGAGCTTCAAT 714  
QY 250 LeuSerGlyValGluIleTyrSerLeuGlnSerSerArgAsnProThrProArgGlySer 269  
DB 715 CTCACCGGCTAGAGATTTACTCGGTTCAATCGTCACGAGAGCCGACGCGGAGAGCTTCT 774  
QY 270 SerPheAsnHisAlaAspPhe-----277  
DB 775 AGCTTTAATCAGACAGATTTCTACGCAATGTTTAAACGAAGCAAGCTCCAAAGCCCTCGT 834  
QY 278 -----AsnIleValGlyAlaAlaAlaLysGlyGly-----288  
DB 835 CACGGTTACACTAATAGTACGCGCGCTGGAGCTGGTCCAGGTGGAGATGTTTACTCA 894  
QY 288 -----288  
DB 895 CTTTCACTCTTAAAGCGGTGACGCGGAGAACGTCAAATTTTGTATGAGGAAGTTATGAG 954  
QY 288 -----288

DB 955 ACGGCGAAGAAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1014  
QY 289 -----GlyAlaAla 291  
DB 1015 AGTGTTCGCTGTCATCCACCGCGCAACCCCAATGTTTACGGGGTCAACGAGTGGAGCAAGT 1074  
QY 292 GlyAspGluGlyLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 311  
DB 1075 GGAGTCAAGAAAGAAAGAGTGGTGGCGGAGGAGCGGTGGC-----1116  
QY 312 AlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTyrSerSerSerAla 331  
DB 1117 GGAAGTAGGATAGGAGGACAAACCAAGGATGATCAATGTTGCTGTGGAGTTCGAGTCT 1176  
QY 332 SerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAla-----348  
DB 1177 TCTCCGTTGTCGAGGACCAACGCGAAGATGCTATGACCAAGAGGTTCTTCTCCACCGATGTA 1236  
QY 349 -----AspHisAlaAspValLeuAlaLysGlyAlaGln 359  
DB 1237 TCCACCGACCTAAAGTTTCTATTCTCTCTCCACGACACCTCTCTCTCTCTCTCTCTCTCT 1296  
QY 360 AlaTyrAspGluTyr-----GlyArgAspAspTyrSerSerArgThrLysAsnGly 376  
DB 1297 AATCTGATAGAACATGTCACCGGAGAAAGGCGATGTGGAATGGACCAAGCGGT 1356  
QY 377 SerGlyGlyAlaAspLysGlyGlyProThrLeuSerLysLeuGlySerAsnSerThrAla 396  
DB 1357 AATAACGGGGAAAG-----TCACCTTACATGGCAGAAAGGTAGCGAC-----1401  
QY 397 GlnLeuTyrProLysAspAspGlyGlyGlyArgAlaAlaValAlaMetProProAla 416  
DB 1402 -----GTGGAAGACGGCGGTCCCGGTCCTAGGAAACAGCAGATGCCCGCGCG 1449  
QY 417 SerValMetThrArgLeuIleLeuIleMetValTyrArgLysLeuIleArgAsnProAsn 436  
DB 1450 AGTGTGATGACGAGACTAATTCGTATAATGTTTGGAGAAACTCATTCGAAACCTTAAC 1509  
QY 437 ThrTyrSerSerLeuIleGlyValValTyrSerLeuValSerTyrArgTyrGlyIleGlu 456  
DB 1510 ACTTACTCTAGTCTTCTTGGCCCTTGTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1569  
QY 457 MetProAlaIleAlaArgSerIleSerIleLeuSerAspAlaGlyLeuGlyMetAla 476  
DB 1570 ATGCCAACGATATGAGTGGATCGATTTCGATATTATCTGATCTGCTGCTTGAATGGCT 1629  
QY 477 MetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLys 496  
DB 1630 ATGTTAGTCTTGTCTATTATGGCATTCACCAAGAAAGATTATCGTGGGAAATCA 1689  
QY 497 LeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAlaAla 516  
DB 1690 GTAGCAGGTTGGGATGGCGGTAAAGTTCTTGACTGGACCGCTGATCGCAGCCACC 1749  
QY 517 SerIleAlaValGlyLeuArgGlyValLeuLeuHisIleAlaIleValGln 533  
DB 1750 TCAATGCAATGTTATTCGAGGTGATCTCTCCATATGCCATCGCTTCAG 1800  
RESULT 14  
US-09-938-842A-1305  
; Sequence 1305, Application US/09938842A  
; Publication No. US20040009476A9  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Krepes, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; FILE REFERENCE: SCRIP1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24

; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 1305  
; TYPE: DNA  
; ORGANISM: Arabidopsis thaliana  
US-09-938-842A-1305

Alignment Scores:  
Pred. No.: 2,96e-136 Length: 1983  
Score: 1406.50 Matches: 319  
Percent Similarity: 59.81% Conservative: 62  
Best Local Similarity: 50.08% Mismatches: 115  
Query Match: 48.33% Indels: 141  
DB: 11 Gaps: 15

US-10-030-884-14 (1-573) x US-09-938-842A-1305 (1-1983)

Qy 1 MettleThrAlaLeuAspLeuTyHisValLeuThrAlaValProLeuTyValAla 20  
Db 1 ATGATCACCGCAAGACATGACGATGTTTACGGCTATGTGCGCTATACGTTGCT 60  
Qy 21 MetThrLeuAlaTyGlySerValArgTrpArgIlePheThrProAspGlnCysSer 40  
Db 61 ATGATATTAGCTATGGTTCGTACGGTGGTGGGGATATTACACCGGACCAATGTTCC 120  
Qy 41 GlyIleAsnArgPheValAlaLeuPheAlaValProLeuLeuSerPheHisPheIleSer 60  
Db 121 GGTATAAACCGGTTGCTTGGCGTTTCGGGTTCCTCTCTCTCTCTCTCTCTCTCTCTCT 180  
Qy 61 ThrAsnAspProPheAlaMetAsnLeuArgPheLeuAlaAlaAspThrLeuGlnLysVal 80  
Db 181 TCCAAATGATCTTTATGCAATGAATTTACCATTCTCTCGCTGCTGATTTCTTTTCAGAAAGTC 240  
Qy 81 AlaValLeuAlaLeuLeu-----AlaLeuAlaSerArgGlyLeuSerSerPro 96  
Db 241 GTTATCTCTCGCGCACTCTTTCTTTGGCAGCGGTTTAGCCGACAGGA----- 288  
Qy 97 ArgAlaLeuGlyLeuAspTrpSerIleThrLeuPheSerLeuSerThrLeuProAsnThr 116  
Db 289 -----AGCTAGAAATGATGATAACGGTCTTTTCACTATCAACACTGCCCTAACACG 339  
Qy 117 LeuValMetGlyIleProLeuLeuArgGlyMetTyGlyAlaSerSerAlaGlyThrLeu 136  
Db 340 TTGGTAATGGGATCCCATTCCTTAGGCGCATGTACCGAGACTTCTCC---GGTAACCTA 396  
Qy 137 MetValGlnValValLeuGlnCysIleleTrpTyThrLeuMetLeuPheLeuPhe 156  
Db 397 ATGGTGCAGATCGTGGTGTTCAGAGCATCATATGTTATATGTTATGTTCTCTCTCTCT 456  
Qy 157 GluTyArgAlaAlaArgAlaLeuValLeuAspGlnPheProAspGlyAlaAlaAlaSer 176  
Db 457 GAGTTCCTGGGGCTTAAGCTTCTCATCTCCGAGCAGTTCCTCCGAG---ACGGCTGGTCA 513  
Qy 177 IleValSerPheArgValAspSerAspValValSerLeu---AlaArgGlyAspValGlu 195  
Db 514 ATTACTTCCTCAGAGTTGACTCTGATGTTATCTCTCTTAATGGCCGGTGAACCCCTCCAG 573  
Qy 196 LeuGluAlaGluProAspGlyValAlaGlyAlaGlyAlaValSerSerArgGlyGlyAsp 215  
Db 574 ACCGATGGCGAGATA-----GGAGAC 594  
Qy 216 AlaGlyArgValArgValThrValArgLysSerThrSerSerArgSerGluAlaAlaCys 235  
Db 595 GACGGAAAGCTACAGTGTGGTTCGAAGATCAAGTGGCGGCTCATCAATGATCTCTTCA 654  
Qy 236 -----SerHisSer-----HisSerGlnThrMetGlnProArgValSerAsn 249

Db 655 TTCAACAATCTCACGGCGGAGGACTTAACCTCTCCATGATAAAGCGCGAGCTTCAAAAT 714  
Qy 250 LeuSerGlyValGluIleTySerLeuGlnSerSerArgAsnProThrProArgGlySer 269  
Db 715 CTCACCGCGGTAGAGATTACTCCGTTCAATCGTCACGAGACCGCGAGAGCTTCT 774  
Qy 270 SerPheAsnHisAlaAspPhePhe----- 277  
Db 775 AGCTTTAATCAGACAGATTCTACGCAATGTTTAAACGCAAGCAAGCTCCAAGCCCTCGT 834  
Qy 278 -----AsnIleValGlyAlaAlaLysGlyGlyGly----- 288  
Db 835 CACGGTTACACTAATAGTACGCGCGCTGGAGTGGTCCAGGTGGAGATGTTTACTCA 894  
Qy 288 ----- 288  
Db 895 CTTCACTCTTCTAAAGGCGTGACGCGAGAACGTCAAATTTTGTATGATGAGAGTTATGAAG 954  
Qy 288 ----- 288  
Db 955 ACGGCGAAGAAAGCAGGAAGAGGAGGAGGAGATGATGATGGGAAATTATACAACTAAT 1014  
Qy 289 -----GlyAlaAla 291  
Db 1015 AGTGTTCCTCTGTACCCACCGCGGAAACCAANTGTTACGGGTCACAGAGTGGAGCAAGT 1074  
Qy 292 GlyAspGluGluLysGlyAlaCysGlyGlyGlyGlyGlyGlyGlyHisSerProGlnProGln 311  
Db 1075 GCGAGTCAAGAAAGAAAGTGTGGCGGAGGAGCGGTGGC----- 1116  
Qy 312 AlaValAlaValProAlaLysArgLysAspLeuHisMetLeuValTrpSerSerAla 331  
Db 1117 GAGTAGGAGTAGGAGGACAAACAGAGAGATGATGTTCTGTGGAGTTCGAGTGT 1176  
Qy 332 SerProValSerGluArgAlaAlaValHisValPheGlyAlaGlyGlyAla----- 348  
Db 1177 TCTCCGGTTCGGAAGCCAAACGCGAAGATGTCATGACCAGAGGTCTTCCCACCGATGA 1236  
Qy 349 -----AspHisAlaAspValLeuAlaLysGlyAlaGln 359  
Db 1237 TCCACCGACCTTAAAGTTTCTATTCTCTCTCAGCAACCTCGCTACTTAACGATGACAG 1296  
Qy 360 AlaTyArgGluTy-----GlyArgAspAspTySerSerArgThrLysAsnGly 376  
Db 1297 AATCTGATAGAACATGTCACGGGAGAAAGGCGATGGAATGACCAAGACGCT 1356  
Qy 377 SerGlyGlyAlaAspLysGlyGlyProThrLeuSerLeuGlySerAsnSerThrAla 396  
Db 1357 AATAACGGGGGAAAG-----TCACCTTACATGGGCAAAAGGTAGCGAC----- 1401  
Qy 397 GlnLeuTyProLysAspAspGlyGluGlyArgAlaAlaAlaValAlaMetProAla 416  
Db 1402 -----GTGGAAGACGGCGGTCCCGTCTAGGAAACAGCAGATGCGCGGCG 1449  
Qy 417 SerValMetThrArgLeuIleLeuMetValTrpArgLysLeuIleArgAsnProAsn 436  
Db 1450 AGTGTGATGACGAGACTAATTTCTGATATGTTTGGAGAAACTCATTCGAAACCCCTAAC 1509  
Qy 437 ThrTySerSerLeuIleGlyValValTrpSerLeuValSerTyArgTrpGlyIleGlu 456  
Db 1510 ACTTACTCTAGTCTTTTGGCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1569  
Qy 457 MetProAlaIleIleAlaArgSerIleLeuSerAspAlaGlyLeuGlyMetAla 476  
Db 1570 ATGCCAACGATATGATGATCGATTTCGATATTATCTGATGCTGCTTGGATGCTGCT 1629  
Qy 477 MetPheSerLeuGlyLeuPheMetAlaLeuGlnProArgIleIleAlaCysGlyAsnLys 496  
Db 1630 ATGTTTAGTCTTGGTCTATTATGTCATTCGAAACCAAGATATTATGCGTGGGAAATCA 1689  
Qy 497 LeuAlaAlaIleAlaMetGlyValArgPheValAlaGlyProAlaValMetAlaAla 516  
Db 1690 GTAGCAGGTTTGGGATGGCGTAAAGTCTTGTAGCTGACCGACCGCTGATCGACCCACC 1749

[illegible]

Db 1481 TTGCCATAACT 1492

Search completed: March 3, 2004, 14:06:12

Job time : 1368 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 3, 2004, 00:50:50 ; Search time 10669 Seconds  
(without alignments)  
11249.120 Million cell updates/sec  
Title: US-10-030-884-13  
Perfect score: 2769  
Sequence: 1 ccacgcgtccgtgcgcct.....aaaaaaaaaaaaaaaaaaag 2769

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*  
1: gb\_ba.\*  
2: gb\_hg.\*  
3: gb\_in.\*  
4: gb\_om.\*  
5: gb\_ov.\*  
6: gb\_pat.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vi.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_ma.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vi.\*  
30: em\_hg\_hum.\*  
31: em\_hg\_inv.\*  
32: em\_hg\_other.\*  
33: em\_hg\_mus.\*  
34: em\_hg\_pln.\*  
35: em\_hg\_rod.\*  
36: em\_hg\_mam.\*  
37: em\_hg\_vrt.\*  
38: em\_sv.\*  
39: em\_hgo\_hum.\*  
40: em\_hgo\_mus.\*  
41: em\_hgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2769	100.0	2769	6	AX046826	AX046826 Sequence
2	910.2	32.9	2233	8	AK103181	AK103181 Oryza sat
3	838.6	30.3	1926	6	AX653844	AX653844 Sequence
4	760	27.4	1686	6	AX652923	AX652923 Sequence
5	716.6	25.9	2126	8	AK063976	AK063976 Oryza sat
6	711	25.7	2470	8	AK103208	AK103208 Oryza sat
7	683	24.7	1884	6	AX653686	AX653686 Sequence
8	671.2	24.2	2358	8	AF056027	AF056027 Oryza sat
9	671.2	24.2	2402	8	AK102343	AK102343 Oryza sat
10	669.6	24.2	2343	8	AK099634	AK099634 Oryza sat
11	632.2	22.8	2457	8	AK101504	AK101504 Oryza sat
12	631.2	22.8	2276	8	AK101191	AK101191 Oryza sat
13	626.8	22.6	2293	6	AX046850	AX046850 Sequence
14	626.8	22.6	2293	8	BT008949	BT008949 Triticum
15	618	22.3	1827	6	AX654121	AX654121 Sequence
16	563	20.3	1845	6	AX652849	AX652849 Sequence
17	546.4	19.7	141025	2	AC123528	AC123528 Oryza sat
18	536.2	19.4	2162	6	AX046832	AX046832 Sequence
19	534.4	19.3	91883	8	CNS08CDY	EX000508 Oryza sat
20	484.8	17.5	1255	8	AK059577	AK059577 Oryza sat
21	470	17.0	170493	2	AP005769	AP005769 Oryza sat
22	452.6	16.3	504	6	AX046824	AX046824 Sequence
23	437.8	15.8	177049	8	AP003725	AP003725 Oryza sat
24	428.4	15.5	140844	2	AP005522	AP005522 Oryza sat
25	425.6	15.4	1426	6	AX046822	AX046822 Sequence
26	409.2	14.8	91870	8	AP004800	AP004800 Oryza sat
27	401	14.5	2297	8	AF515435	AF515435 Populus t
28	394.6	14.3	127130	2	AC098571	AC098571 Oryza sat
29	383	13.8	2316	8	BTU249298	AJ249298 Brassica
30	361.4	13.1	2254	8	AF087016	AF087016 Arabidops
31	352.2	12.7	2358	8	BT008691	BT008691 Arabidops
32	352.2	12.7	2404	8	AJ091009	AJ091009 Arabidops
33	351	12.7	2462	8	BJU249297	AJ249297 Brassica
34	348.6	12.6	2581	8	BJU132363	AJ132363 Brassica
35	331.4	12.0	1869	8	AY093960	AY093960 Arabidops
36	331.4	12.0	2261	8	AF372950	AF372950 Arabidops
37	329.8	11.9	2276	8	AF089084	AF089084 Arabidops
38	329.8	11.9	2292	6	A68597	A68597 Sequence 11
39	329.8	11.9	2292	6	A79382	A79382 Sequence 11
40	323.8	11.7	1088	6	AX046816	AX046816 Sequence
41	319	11.5	1860	6	AX506153	AX506153 Sequence
42	319	11.5	1860	6	AX651788	AX651788 Sequence
43	319	11.5	2071	8	AF087820	AF087820 Arabidops
44	316.6	11.4	2549	6	AX046842	AX046842 Sequence
45	305.6	11.0	1796	8	AF087819	AF087819 Arabidops

ALIGNMENTS

RESULT 1  
AX046826  
LOCUS  
DEFINITION Sequence 13 from Patent WO0068389.  
ACCESSION AX046826  
VERSION AX046826.1 GI:11876320  
KEYWORDS  
SOURCE Zea mays  
ORGANISM Zea mays  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.  
REFERENCE  
AUTHORS Orozco, E.M., Weng, Z., Bruce, W.B., Cahoon, R.E. and Tao, Y.  
TITLE Auxin transport proteins

Pred. No. is the number of results predicted by chance to have a









Syngenta Participations AG (CH)  
 Location/Qualifiers  
 1. 1686  
 /organism="Oryza sativa"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:4530"

ORIGIN  
 Query Match 27.4%; Score 760; DB 6; Length 1686;  
 Best Local Similarity 74.4%; Pred. No. 2, le-76;  
 Matches 1198; Conservative 1; Mismatches 256; Indels 155; Gaps 13;

171 ATGATCACCAGCGCTGACCTCTTACACAGTGTGAGCGGGTGTGTCGCCGTGACGTGCC 230  
 Db 1 ATGATCACCAGCGCTGACCTCTTACACAGTGTGAGCGGGTGTGTCGCCGTGACGTGCC 60  
 231 ATGAGCTCGCTGACGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 290  
 Db 61 ATGACGCTGGGTGACGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
 291 GGGATCAACCGCTTCTGTCGCTCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 350  
 Db 121 GGCATCAACCGCTTCTGTCGCTCTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
 351 ACCAAGACCCCTTCCCATGACCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410  
 Db 181 ACCAAGACCCCTTCCCATGACCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240  
 411 GCGCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 470  
 Db 241 ATGCTCTCGCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 291  
 471 CTCGACTGAGCATACGCTCTTCT 529  
 Db 292 CTCGACTGCTCATCAACCT 351  
 530 CATCCCGCTGCTGCGAGGATGATGACGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCTGCTGCGCGCT 589  
 Db 352 CATCCCGCTGCTCAAGGGAT-----CTGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 372  
 590 CGTCTCTCTCGAGTGATCATCTGTGTGATACGCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 649  
 Db 373 -----TGCATCATCTGTGTACACGCTCATCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 419  
 650 CGCGCGCGCTGCTCTCTGACAGTTCCTGACG 709  
 Db 420 CGCGCGCTCTCTCTCATGAGCAGTTCGAGCA---CGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 476  
 710 CGCGCTCGACTCCGACGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769  
 Db 477 CGCGGTGGAATCCGAGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 526  
 770 CGAGCGGCTCG 829  
 Db 527 ---CGAGCTGCGAGCGCGCGCGAG-----GTCGGGACGATGCGCAATGCG 572  
 830 CGTCACCGTCCGAAGTCCACACAGCTCGCGCTCCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 881  
 Db 573 AGTCACCGTCCGAAGTCCGACAGCTCGCGCTCCGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 632  
 882 -TGCACCTCCACACCATGACCGCGCT 940  
 Db 633 GTCACTGCT 692  
 941 GCTGCACTGCT 1000  
 Db 693 GCTGCACTGCT 752  
 1001 CAACATGCT 1060  
 Db 753 CAACATGCT 794  
 1061 CGCATGCG 1120

Db 795 CGC-----CGCGCGTGGCGGGGCGCACTCGCGCGACCG-----GTGGTGG 836  
 Qy 1121 CAAGAGGAGGACCTGCACATGCTCTGCTGGAGCTCCAGCGCTCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1180  
 Db 837 GAAGAGGAGGACCTGCACATGCTCTGCTGGAGCTCCAGCGCTCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 896  
 Qy 1181 CGCGCGCTGTCACGTCTTTCGGCGCGCGCGCTGACCATGCGGACGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1240  
 Db 897 CGCGCGCGCGCGCGCGCGCTCGCGCGCTGACCATGCGGACGTCTCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 956  
 Qy 1241 AGCCCGGCTTACGACGAGTACGCGCGCGCTGACCATGCGGACGTCTCTGCGGACGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1300  
 Db 957 CGCGCGCGCGCGCGCGCGCTGCTGCGCGCTGACCATGCGGACGTCTCTGCGGACGTCTCTGCGGACGTCTCTGCTGCTGCTGCTGCTGCTGCT 1010  
 Qy 1301 CGCGCGCGCGCGCGCGCGCTGCTGCGCGCTGACCATGCGGACGTCTCTGCGGACGTCTCTGCGGACGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1360  
 Db 1011 -----TGAGAGGACGCGCGCGCTGACCATGCGGACGTCTCTGCGGACGTCTCTGCGGACGTCTCTGCGGACGTCTCTGCTGCTGCTGCTGCTGCTGCT 1061  
 Qy 1361 GCTGTACCCCAAGGACGAGCGCGCGCGCTGACCATGCGGACGTCTCTGCGGACGTCTCTGCGGACGTCTCTGCGGACGTCTCTGCTGCTGCTGCTGCTGCTGCT 1420  
 Db 1062 GCTCGCGCGCGCGCGCGCGCTGACCATGCGGACGTCTCTGCGGACGTCTCTGCGGACGTCTCTGCGGACGTCTCTGCGGACGTCTCTGCTGCTGCTGCTGCTGCTGCT 1115  
 Qy 1421 CGTGATGACCGCGCTCATCTCTCATGCT 1480  
 Db 1116 CGTGATGACGAGGCTCATCTGATGCT 1175  
 Qy 1481 CTACTCAGCGCTCATCGCGCT 1540  
 Db 1176 TTATCCAGCGCTCTCGGTGCT 1235  
 Qy 1541 GCCAGCGCATCATCGCGCGCT 1600  
 Db 1236 GCGCGCTATCATCGCGCGCT 1295  
 Qy 1601 GTTCAGCGCTAGGCT 1660  
 Db 1296 GTTCAGCGCTAGGCT 1355  
 Qy 1661 GCGCGCGCATCATCGCGCGCT 1720  
 Db 1356 TGCTTCTGATGCT 1415  
 Qy 1721 CATCGCGCT 1770  
 Db 1416 CATCGCGCT 1465

## RESULT 5

AK063976  
 LOCUS  
 DEFINITION  
 Oryza sativa (japonica cultivar-group) cDNA clone:001-124-C02, full insert sequence.

## ACCESSION

AK063976

## VERSION

AK063976.1 GI:32973994

## KEYWORDS

FLI\_CDNA, oligo-capping.

## SOURCE

Oryza sativa (japonica cultivar-group)

## ORGANISM

Oryza sativa (japonica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1

The Rice Full-length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-length cDNA Project Team,

Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Hotta, I.,

Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,

Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,

Ohtsuki, K., Shishiki, T., Foundation of Advancement of International

Science Genome Sequencing & Analysis Group; Otomo, Y., Murakami, K.,

Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Teunoda, Y.,

Kurosaki, R., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,

Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J.,

Ikedo, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,

Kusumegi, T., Oka, M., Ryu, R., Deda, M., Matsubara, K., RIKEN;  
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Oeato, N., Oka, Y.,  
Saito, E., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Yoshino, M., and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from  
japonica rice

Science 301 (5631), 376-379 (2003)  
22752273  
12869784

2 (bases 1 to 2126)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,  
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, Y., Ikeda, R., Imamura, K.,  
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
Kagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
Namiki, T., Narikawa, R., Niikura, J., Nishii, K., Nomura, K.,  
Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
Osato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
Toya, T., Tsunoda, Y., Ueda, X., Waki, K., Xie, Q., Yahagi, W.,  
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and  
Yoshimura, A.

Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of  
Agrobiological Sciences, Department of Molecular Genetics, Head of  
Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  
305-8602, Japan [E-mail:sk.kuchin@nias.affrc.go.jp,  
Tel:81-29-858-7007, Fax:81-29-858-7007]

This clone is one of the 28K full-length cDNA clones from japonica  
rice.  
URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,  
Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and  
Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,  
Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,  
Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
Yoshimura, A., Matsubara, K. and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center  
and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kouda, M.,  
Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oeato, N.,  
Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers  
1. .2126  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol type="mRNA"  
/culti var="Nipponbare"  
/db\_xref="taxon:39947"

TITLE  
JOURNAL

COMMENT

FEATURES  
source

ORIGIN	/clone="001-124-C02"
Query Match	25.9%; Score 716.6; DB 8; Length 2126;
Best Local Similarity	85.8%; Pred. No. 1.5e-71;
Matches 1125; Conservative	0; Mismatches 554; Indels 36; Gaps 4;
Qy	163 CCGCGGCGATGATCAACGGCGCTGGACCTCTACACGCTGTGACGGCGGTGGTCCCGCTGT 222
Db	100 CTGAGGAGATGATATCCGGGCACGACTTCTACACGGTGTATGGCGGGTGGTCCCGCTGT 159
Qy	223 ACCTGGCCATGACGCTGGGGTACGGCTCCGTCGGCTGGTGGCGGCATCTTCACGCCGAGACC 282
Db	160 ACCTGGCGATGTTTCTTGGGGTACGGGTGGTGGGGTGGTGGGGCATCTTCACGCCGAGACC 219
Qy	283 AGTGTCTCCGGGATCAACCGCTTTCGTCGGCGCTCTTCGCCGCTGCGCGCTCTCTCTCCACT 342
Db	220 AGTGTCTCCGGCATCAACCGCTTTCGTCGGCATCTTCGCCGCTGCGCGCTCTCTCTCCACT 279
Qy	343 TCATCTCCACCAACGACCCCTTGGCCATGAACCTGGGCTTCTTGGCGGGCGGACACGCTGC 402
Db	280 TCATCTCCACCAACGACCCCTTGGCCATGAACCTTCCGCTTCTTGGGGGGGGACACGCTGC 339
Qy	403 AGAAGTGGCGGCTCTCGCGCTGCTGGCGCTGCGCTCCCGCGGCGCTCTCTCCCGCGCGG 462
Db	340 AGAAGTGGCTGCTCTGGGGGGCTGCGCGCTGCTGGCGCTCTCTCCCGCGGCGGCG 399
Qy	463 CGCTCGGGCTCGAGCTGGAGCATCAGCGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 522
Db	400 CGCGCGGCTGAGCTGGTCCATCACGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 459
Qy	523 TCATGGGCATCCCGCTGCTCGAGGATGATGACGGCGCTGCTGGCGGGCGGACGCTCATGG 582
Db	460 TCATGGGGATCCCGCTGCTGATGCCATGATACGG---GCCATCTCCGGCTCTCTCTCTCT 516
Qy	583 TCCAGGTGCTGCT 642
Db	517 TCCAGATGCTGCT 576
Qy	643 ACCGCGCGCGCGCGCGCTGCT 702
Db	577 TCCGCGCGCGCGCGGATGCTGATCGCGACAGTTCCTCGGA---CAGCGCGCGCTCTCTCT 633
Qy	703 TCTCTCTTCCCGCTGCT 762
Db	634 TGTCCTGTCAGCTGACCGCGAGCTGCTGCTGCTGAGGGCGCGCCACCGCGAGAGCGAGG 693
Qy	763 CCGAGCCCGACGGGCTGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 822
Db	694 CCGAGGTGGCGCGGACGGCGGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 753
Qy	823 GGGTGGCGCTCACCGTGGCGAAGTCCACAGCTCTCGGCTCCGAGGCGCGCTGCTCTCTCT 882
Db	754 GGTCTGCTGCTGCT 813
Qy	883 CGCACTCCAGACCATGACCGCGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 942
Db	814 GCTCGTCCGGAACCCAAACCGCGCGGCTCTCAACTTCAACCCAGCGGCTCTCTCTCTCT 873
Qy	943 TGCAGTCTGCTCGCGCAACCCCGCGCGGCTCTCACTTCAACCCAGCGGCTCTCTCTCTCT 1002
Db	874 TGGTGGCGCGCGGCGGCT 933
Qy	1003 ACATCTGTCGGCGCGCGCGCGCAAGGAGGCGGAGGAGGCGGGGGGACGAGGAGAGGCGG 1062
Db	934 CCTCCGAGCTCTACTCTGCTCAATCTGCTGCGGGGGCGCGCAACCCCGAGGCGAGTCTCA 993
Qy	1063 CATCGCGCGCGCGGCGGAGGACATCGCGCGGAGCGGCGGCTGCTCTCTCTCTCTCTCT 1122
Db	994 ACGAGCTCTCGGCGCGCGCGCGCGGAGGACCGCGGAGCGGCGGCTCTCTCTCTCTCTCT 1053
Qy	1123 AGAGGAGGAGCTGCGCATCTCTGCTGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1182

Db 1054 ATGCCAAGAGCTCCACATGTTCTGTTGAGCTCGAGCGCGTCTCCGCTCTCAGAAAGTCA 1113

Qy 1183 CGCCCGTGCACGTTCTTGGCGCGCGCGGCGTGAACATCCGACGCTCTCTCGCCAAAGAG 1242

Db 1114 GCGGCTGCTGTTTCACTGCGCGCGCGCGCGCGTCTCGACGTCGCGCGCAAGGAAA 1173

Qy 1243 CCA---GCGTACGACGAGTACGCGCGCGCGTACGACGACGAGGAGGAAACGCGA 1299

Db 1174 TCACATGTTCTATCCCGCGAGCTCCCGCAGAACACGCGTACGCGAAGACGACGAG 1233

Qy 1300 GCGCGCGCGGCAAGAGCGCGCGCGACGCTGTGCAAGCTGGGGTCCAACTCGACGCGCG 1359

Db 1234 AGTACGCGCGAGTGGCATTTGGGTGGCGCGCGCGCGGAGAGAACTTCAGCTCGACGCGCG 1293

Qy 1360 AGCTGTACCCAGGACGA-----CGCGAGGCGGAGG 1392

Db 1294 AGTGTACCCAGGAGTGTGTCGAGTTCGACGACGACCGAACCGCGCGCGCGCGCGCGCG 1353

Qy 1393 CGCGCGCGTGGCGATCCCGCGCGCGGAGGATGATGACGCGGCTCATCTCTCATCATGTTGT 1452

Db 1354 CGCGCGAGTACCAATATCGCGCGCGGAGGATGATGACACGCTCATCTCTCATTAATGTTGT 1413

Qy 1453 GAGGAGAGTATCCGACACCCCAACACCTACTTCAGGCTCATCGCGCTGCTCTGTCGCC 1512

Db 1414 GCGCAAGTCTATCCCAACCCCAACACTTATCCAGGCTCTCTCGGCTTCGCTGCTGCC 1473

Qy 1513 TGTCTCTACAGTGGGCGATCGAGATGCCAGCGATCATCGCGCGTTCGATTCGATCC 1572

Db 1474 TCGTCGCTTCGCTGCGAGTCTCATGCGAGCAATCGTCGAGAGTCCATCTCCATTC 1533

Qy 1573 TGTGCGACGCGGCTCTCGGAGTGGCCATGTTTCAGCTTAGGCTGTTTCATGCGCGCTGCGAGC 1632

Db 1534 TCTCGGACGCGAGGCTCTGGGAGTGGCCATGTTTAGCTGGGATGTTTCATGCGCGCTGCGAGC 1593

Qy 1633 CGAGGATCATCGGTCGCGGACCAAGCTGCGGCGCATCGGATGCGGCTGCGGTTTCGTCG 1692

Db 1594 CAGCATCATCGGCTGTGCAATACGCGCGCTGCTCTCATGCGCTTCGCTTCCTCG 1653

Qy 1693 CAGGCGCGCGGTCATGCGCGCGCTTCATCGCGCTGCGGTCGCTGCGCGCGCTTCCTCC 1752

Db 1654 CCGGCGCTGCGGTCATGCGCGCGCTCAATCGGCTGCGGATGTTTCATGCGCGCTGCGAGC 1713

Qy 1753 ACATCGGCTGCTGCGGAGTCTGCTGCTGCGGATGCTGCGGTCGCTGCTGCTGCTGCTGCTG 1812

Db 1714 ACCTGCGCATGTTTCAGGCGGCTCTACCAAGGAGTGTGCTTGTGTTTTCGAAAG 1773

Qy 1813 AGTACGCGGTTTCATCCCGCATCTTCGAGCAGCG 1847

Db 1774 AATACATGTCACCGCGCGCATCTCGAGCAGCG 1808

# RESULT 6

AK103208  
LOCUS  
DEFINITION  
Oryza sativa (japonica cultivar-group) cDNA clone:J033122123, full insert sequence.

ACCESSION  
AK103208  
VERSION  
AK103208.1 GI:32988417  
KEYWORDS  
FLI\_CDNA; CAP trapper.  
SOURCE  
Oryza sativa (japonica cultivar-group)  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzae; Oryza.

REFERENCE  
1  
The Rice Full-length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-length cDNA Project Team; Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,

Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Nariawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Iehibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, A., KIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ootomo, Y., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M. and Hayashizaki, Y.  
Collection, mapping, and annotation of over 28,000 cDNA clones from japonica rice  
Science 301 (5631), 376-379 (2003)  
22752273  
PUBMED  
12869764  
2 (bases 1 to 2470)  
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashizaki, Y., Hayatsu, N., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imotani, K., Iehibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kikuchi, S., Kishikawa-Hirozane, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Nariawa, R., Niikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohtsuki, K., Oka, M., Ooka, H., Ootomo, Y., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akanishi, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.  
Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
This clone is one of the 28K full-length cDNA clones from japonica rice.  
URL: http://cdna01.dna.affrc.go.jp/cdna/  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.  
FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Iehibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Nariawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kouda, M., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohtsuki, N., Ootomo, Y., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akanishi, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.  
Location/Qualifiers  
1..2470  
/organism="Oryza sativa (japonica cultivar-group)"

## FEATURES

source

```
/mol_type="mRNA"  
/cultivar="Nipponbare"  
/db_xref="taxon:39947"  
/clone="J033122I23"
```

## ORIGIN

Query Match 25.7%; Score 711; DB 8; Length 2470;  
Best Local Similarity 67.5%; Pred. No. 6.3e-71;  
Matches 1193; Conservative 0; Mismatches 430; Indels 144; Gaps 8;

RESULT 7	REFERENCE
AX653686	AUTHORS
LOCUS	
DEFINITION	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	

AX653686 1884 bp DNA linear PAT 22-MAR-2003  
Sequence 3556 from Patent WO0300898.  
AX653686  
AX653686.1 GI:29156500  
.  
Oryza sativa  
Oryza sativa  
Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzaceae; Oryza.  
1  
Kang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M.,  
Chang, H. F., Quan, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.  
2003



TITLE BIR1, a root-specific protein involved in auxin transport, is required for gravitropism in Arabidopsis thaliana  
 JOURNAL Genes Dev. 12 (14), 2175-2187 (1998)  
 MEDLINE 98344010  
 PUBLISHED 9679062  
 REFERENCE 2 (bases 1 to 2358)  
 AUTHORS Luschnig, C., Grisafi, P. and Fink, G.R.  
 TITLE Direct Submission  
 JOURNAL Submitted (30-MAR-1998) Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, MA 02142, USA  
 FEATURES  
 Location/Qualifiers  
 1..2358  
 /organism="Oryza sativa"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4530"  
 1..2358  
 /gene="REH1"  
 /size=152..1939  
 /gene="REH1"  
 /note="potential membrane protein"  
 /codon\_start=1  
 /product="auxin transport protein REH1"  
 /protein\_id="AAC39514.1"  
 /db\_xref="GI:3377509"  
 /translation="MITAADFVHVTAMVPLVYVAMI LAYGVKWRIFTPDQSGINR FVALFVLLGSHFISTNNPVTMLRFLAADTLOKLVMLAMLTAWSHLSRSGLEWITLFSLTPLNPTLVMGIVPILKMGYGFSGSLMVQIVLQCIWYITLMLFMPEYRGRMLITEQPDTAANATISVDPDVVSDGRDALETETEVKEDGRIVHTVRRSNASRDIVSRSSNGFSSTPRPNLNIAIYLSLOSRENTPRGSFNFHTDFYSMVGRSNGAADA FGVRTGATPRSNYEDDASKPKYIPLFASNAFPMAGHPYAPNPDAVSSAPKGAATNG QAKGDELHMFVSSASVSDVDFGGAPDYNDAAVSPKRGDGAQKREDYVERDDFS FGNRGMDRDAEAGDEKAAAGADPSKMAAAPTAMPTSYMTLILIMVWRKLIRNP NTYSSLLIGLIVCFRNWFEMPAIVLKSISLDAGLGMAFSLIGLIFMALQPHIAC GNKVATYAMVRFIAGPAVMAAASPAVLGRGLTHLVAIQVQALPQGIIVPFVFAKEYSV HPSILSTAVIFGMLIALPITLVYIILGL"  
 ORIGIN  
 Query Match 24.2%; Score 671.2; DB 8; Length 2358;  
 Best Local Similarity 66.8%; Pred. No. 1.9e-66;  
 Matches 1196; Conservative 0; Mismatches 423; Indels 171; Gaps 10;  
 QY 165 GCGGCATGATCACCGCGTGGACCTTACCACGCTGACGGCGGTGGTCCGCGTGTAC 224  
 DB |||||  
 QY 146 GCGAGATGATTACGGCGCGGACTTCTACACGCTGATGACGGGATGGTCCGTTGAC 205  
 DB |||||  
 QY 225 GTGGCCATGACGCTGGCGTACCGGCTCCGCTCCGCTGGTGGCGCATCTTACGCGGACCG 284  
 DB |||||  
 QY 206 GTGGCGATGATCTGGCGTACGGGTCGGTGAAGTGTGGGCGATCTTCAAGCGCGACCG 265  
 DB |||||  
 QY 285 TGCTCGGGATCAACCGCTTGTGGGCTCTTCGCGGCTGGCGCTCTCTCTCTTCACATTC 344  
 DB |||||  
 QY 266 TGCTCGGGATCAACCGCTTGTGGGCTCTTCGCGGCTGGCGCTCTCTCTCTTCACATTC 325  
 DB |||||  
 QY 345 ATCTCCACCAACGACCCCTTCGCCATGAACCTGCGCTTCTTCCGCGCGGACACGCTGCAG 404  
 DB |||||  
 QY 326 ATCTCCACCAACGACCGTACAGTGAACCTTCGGTTTCACTCGCGCGGACACGCTGCAG 385  
 DB |||||  
 QY 405 AAGGTGGCGCTCTCGCGCTGTGGGCTGGGCTCCGCGGCGCTCTCTCTCTCTCTCTCTCT 464  
 DB |||||  
 QY 386 AAGCTGATGGTGTGGCCATGTCTACGGGCTGGAGCCACCTCAGC-----CGCGG 436  
 DB |||||  
 QY 465 CTCGGGCTCGACTGGAGCATACGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 524  
 DB |||||  
 QY 437 GGGAGCTCGAGTGGACCATCAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 496  
 DB |||||  
 QY 525 ATGGGCATCCCGCTGTGCGAGGCATGTACGCGCGCTGTGCGCGGACGCGCTCATGCTC 584  
 DB |||||  
 QY 497 ATGGGATCCCTTTGTCTAAGGCGATGTACGGGAGTTCT---CGGCGAGCGCTCATGGTG 553  
 DB |||||  
 QY 585 CAGGTGCTGTCTCCAGTGCATCTCTGGTACAGCTCATGCTCTCTCTCTCTCTCTCTCTCT 644  
 DB |||||  
 QY 554 CAGATGCTGTGTGAGTGCATCTCTGTGTACAGCTCATGCTCTCTCTCTCTCTCTCTCTCT 613

QY 645 CGCGCGCGCGCGCGCTCGTCTCGACCACTTCCCGACGGCGCGCGCGCTCCATCGTC 704  
 DB |||||  
 QY 614 CGCGCGCGCGGATGTCTATCACCGAGCATTTCCCGA---CACCGCGCGCAATCGCC 670  
 DB |||||  
 QY 705 TCCTTCGCGCTCGACTCCGACGTCGCTCGCTCGCGCGGAGACGCTCGAGCTCGAGGCC 764  
 DB |||||  
 QY 671 TCCATCGTGTGACCGGAGCTGCTGCTGGACGGCAGGAGGACGCCATCGAGACG 730  
 DB |||||  
 QY 765 GAGCCGACGGCTCGCGCGCGCGCGCGCTCTCTCTCGCGCGGAGACGCGCGCGCG 824  
 DB |||||  
 QY 731 GAGACGGA-----GGTGAAGGAGGACGGCAGG 757  
 DB |||||  
 QY 825 GTGCGGCTCACCGTGGCGCAAGTCCACAGCTCGCGCTCGAGCGCGCGCTCGCCTC- 883  
 DB |||||  
 QY 758 ATACAGTCACTCGTGGCGCTCCAAAGCGCTCTCGCTCGGACATCTACTCCCGCGCTCC 817  
 DB |||||  
 QY 884 --GCATCCGACACATGACAGCCCGCTGTGTCGAACCTCTCGCGCTGAGATCTACTCG 941  
 DB |||||  
 QY 818 ATGGGCTTCTCCAGCACCAAGCGCGCGCGGAGCAACTCACCACCGCGAGATCTACTCG 877  
 DB |||||  
 QY 942 CTGCACTGCTGCGGCAACCCACCGCGCGGCTCCAGCTTCAACCAACGCGGACTTCTTC 1001  
 DB |||||  
 QY 878 CTGCACTGCTGCGGCAACCGCGCGCGGCTTCAAGCTTCAACCAACGCGGACTTCTAC 937  
 DB |||||  
 QY 1002 AACATGCTGCGCGCGCGCGCAAGGAGCGGAGGAGCGCGCGGAGACGAGGAGAGGGC 1061  
 DB |||||  
 QY 938 TCCATGCTGGCGCGGCTCCAACTTCGCGCGCGCGCGGCTTCGCGCTCCGACCGCGC 997  
 DB |||||  
 QY 1062 GC-----CGTCCG 1063  
 DB |||||  
 QY 998 GCCACGCGCGCGCGCTCCAACTAGAGGAGCGCGCTCCAGCCCAAGTACCGGCTCCG 1057  
 DB |||||  
 QY 1064 -----ATGCGCGCGCGCGCGCGGAGAC-ACTGCGCGAGCGCGCGCGCGCGCGCG 1111  
 DB |||||  
 QY 1058 GCGTCAATGCG 1117  
 DB |||||  
 QY 1112 -----CGTCCG 1139  
 DB |||||  
 QY 1118 GCGCCCAAGCG 1177  
 DB |||||  
 QY 1140 ATGCTGCTGTGGAGCTCCAGCGCTCGCGCGCTGTCGAGCGCGCGCGCGCGCGCGCG 1199  
 DB |||||  
 QY 1178 ATGCTGCTGTGGAGCTCCAGCGCTCGCGCGCTGTCGCGCGCGCGCGCGCGCGCG 1222  
 DB |||||  
 QY 1200 GCGCGCGCGCGCGCTGACATGCGCGAGCTCTCGCGCGCGCGCGCGCGCGCGCGCG 1259  
 DB |||||  
 QY 1223 GCGCGCGCGCGCGCGAGACTACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1282  
 DB |||||  
 QY 1260 TACGCGCGCGCGCGCTACAGCAGCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1319  
 DB |||||  
 QY 1283 GAGCGGAGGACAGGAGGAGCTAGCTGGAGCGGAGCGATTTCACTTCGGGACAGGGG 1342  
 DB |||||  
 QY 1320 G---GGCGCGCGCTGTGAGCTGGGCTCCAACTCGACGGCGCGCTGTATCCCAAGGAC 1376  
 DB |||||  
 QY 1343 GTCATGACAGGACCGCGGAGGAGCGGAGCGAGAGCGCGCGCGCGCGCGCGCGCGCG 1402  
 DB |||||  
 QY 1377 GAGCGCGGAGGAGCGCGCGCGGCGGATCCCGCGCGAGCGCTGATGACGCGCGCTC 1436  
 DB |||||  
 QY 1403 CCCAGCAAGGCG 1462  
 DB |||||  
 QY 1437 ATCTCATCATGTGTGGAGGAGCTGATCCGGAACCCCAACACCTTACTCCAGCTCATC 1496  
 DB |||||  
 QY 1463 ATCTGATCATGTGTGGCGAAGCTCATCCGCAACCCGCAACCTTACTCCAGCTCATC 1522  
 DB |||||  
 QY 1497 GCGCTGCTGTGTCCTGTCTCTTACAGTGGGCGATGAGATGCGCGCGCGCGCGCGCG 1556  
 DB |||||  
 QY 1523 GCGCTCATGTGTGTCCTGTCTCTTACAGTGGGAACTTCAGATGCGCGCGCGCGCGCG 1582  
 DB |||||  
 QY 1557 CGGTTCATTTTCGATCTGTGCGAGCGGCGGTCTCGGATGCGCGCGCGCGCGCGCGCG 1616  
 DB |||||  
 QY 1583 AATTCATCTCGATCTGTGCGAGCGCGGCGCTCGGATGCGCGCGCGCGCGCGCGCG 1642  
 DB |||||  
 QY 1617 TTCATGCGCGCTGACGCGGAGGATCATCGCTGCGGCGGNAACAGCTGCGCGCGCGCG 1676



Db 496 ATGGGGATCCCTTTCTCAAGGCGCATGTACGGGAGTTCT---CGGCGAGCTCATGGTG 552  
 Qy 585 CAGGTCTGTCCTCCAGTGCATCTGTTGATACAGCTCATGCTCTTCTTCTTCCAGTAC 644  
 Db 553 CAGATCTGTCGTCTGATGATCATCTGTTGATACAGCTCATGCTCTTCTTCTTCCAGTAC 612  
 Qy 645 CGCGCGCGCGCGCTCTCTCTCCAGCAGTTTCCCGACGCGCGCGCGCTCTCTCTCTCTCT 704  
 Db 613 CGCGCGCGCGCGATCTCATCACCGAGCAGTTTCCCGGA---CACCGCGCGCAACATCGCC 669  
 Qy 705 TCCTTCGCGTGCATCTCGAAGTCTGCTGCTGCGCGAGGGGAGAGTGCAGCTCAGAGCC 764  
 Db 670 TCCATCTGTCGTGACCCCGGAGTCTGCTGCTGAGCGGAGGAGAGCCATCGAGACG 729  
 Qy 765 GAGCCCGGAGCGGCTCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 824  
 Db 730 GAGACGGA-----GGTGAAGGAGAGCGGAGG 756  
 Qy 825 GTGCGCGTACCGTGGCAAGTCCACAGCTCGGCTCCGAGGCGCGGTGTCGCACTC- 883  
 Db 757 ATACACGTCACCGTGGCGGCTCCAAACGCGTCTGCTCGGACATCTACTCCCGCGCTCC 816  
 Qy 884 --GCACCTCCAGACCATGACGCGCGCTGTGTCACCTCTCTCGGCTGGAGATCTACTCG 941  
 Db 817 ATGGCTTTCTCAGACCAACGCGCGCGCGGAGACCTCAACACCGCGAGATCTACTCG 876  
 Qy 942 CTGCACTGTCGCGCAACCCCAACCGCGCGCGGTCAGCTTCAACCAACCGCGAGCTTCTTC 1001  
 Db 877 CTGCACTGTCGCGCAACCCCAACCGCGCGGTTCAAGCTTCAACCAACCGCGAGCTTCTTC 936  
 Qy 1002 AACATCTGCGCGCGCGCGCGCGGAGGCGGAGGCGGCGGCGGCGGAGGAGGCGGCGG 1061  
 Db 937 TCCATGGTGGCGCGAGCTTCAACTTCGCGCGCGCGCGGCTTGGCGGCTCCGACCGCG 996  
 Qy 1062 GC----- 1063  
 Db 997 GCACGCGCGCGCGCTTCAACTACAGGAGCGCGGTCACAGCCCAAGTACCGGCTCCG 1056  
 Qy 1064 -----ATGCGCGCGCGCGCGGAGGAC-ATGCGCGCGCGCGCGCGCGCGCGCGCG 1111  
 Db 1057 GGTGCAATGCG 1116  
 Qy 1112 -----CGTGGCG 1139  
 Db 1117 GCGCCCAAGGCGCGCAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1176  
 Qy 1140 ATGCTCTGTCGAGTCCAGCGCTCGCGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCG 1199  
 Db 1177 ATGTTCTGTCGAGTCCAGCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1221  
 Qy 1200 GCGCGCGCGCGCGTGCATCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1259  
 Db 1222 GCG 1281  
 Qy 1260 TACG 1319  
 Db 1282 GCG 1341  
 Qy 1320 G---GGCCG 1376  
 Db 1342 GTCATGGACG 1401  
 Qy 1377 GACG 1436  
 Db 1402 CCGAGCAAGCG 1461  
 Qy 1437 ATCTCATCATGCTGTGAGGAGTGTATCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1496  
 Db 1462 ATCTCATCATGCTGTGAGGAGTGTATCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1521  
 Qy 1497 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1556

Db 1522 GGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1581  
 Qy 1557 CGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1616  
 Db 1582 AATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1641  
 Qy 1617 TTTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1676  
 Db 1642 TTTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1701  
 Qy 1677 GCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1736  
 Db 1702 GCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1761  
 Qy 1737 CGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1796  
 Db 1762 CGTGGCACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1821  
 Qy 1797 TTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1846  
 Db 1822 TTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1871

# RESULT 10 AK099634

LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J013059F15, full  
 DEFINITION insert sequence.

ACCESSION AK099634  
 VERSION AK099634.1 GI:32984843

KEYWORDS FLJ CDNA; CAP trapper.

SOURCE Oryza sativa (japonica cultivar-group)

ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzeae; Oryza.

## REFERENCE

### AUTHORS

The Rice Full-length cDNA Consortium, National Institute of  
 Agrobiological Sciences Rice Full-length cDNA Project Team,  
 Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,  
 Kojima,M., Yamaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,  
 Kishimoto,N., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,  
 Ohtsuki,K., Shishiki,T., Foundation of Advancement of International  
 Science Genome Sequencing & Analysis Group, Ohtsuki,Y., Murakami,K.,  
 Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,  
 Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,  
 Nariawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Niikura,J.,  
 Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,  
 Kusumegi,T., Oka,M., Ryu,R., Ueda,M., Matsubara,K., RIKEN,  
 Kawai,J., Carninci,P., Adachi,J., Aizawa,K., Arakawa,T., Fukuda,S.,  
 Hara,A., Hashidume,W., Hayatsu,N., Imotani,K., Ishii,Y., Itoh,M.,  
 Kagawa,I., Kondo,S., Konno,H., Miyazaki,A., Osato,N., Ota,Y.,  
 Saito,R., Sasaki,D., Sato,K., Shibata,K., Shinagawa,A., Shiraki,T.,  
 Yoshino,M. and Hayashizaki,Y.

Collection, mapping, and annotation of over 28,000 cDNA clones from  
 Japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)

MEDLINE 22752273

PUBMED 12869764

REFERENCE 2 (bases 1 to 2343)

### AUTHORS

Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K.,  
 Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W.,  
 Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiraoka,T.,  
 Hori,F., Hotta,I., Iida,J., Ikeda,R., Imamura,K.,  
 Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I.,  
 Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M.,  
 Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,  
 Koya,S., Kurihara,C., Kusumegi,T., Kusumegi,T., Li,C., Lu,M.,  
 Masuda,H., Matsubara,K., Matsuyama,T., Miura,J., Miyazaki,A.,  
 Mizuno,K., Murakami,K., Murata,M., Negata,T., Nakamura,M.,  
 Namiki,T., Nariawa,R., Niikura,J., Nishi,K., Nomura,K.,  
 Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H.,

Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.  
Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel: 81-29-838-7007, Fax: 81-29-838-7007)  
This clone is one of the 28K full-length cDNA clones from japonica rice.  
URL: <http://cdna01.dna.affrc.go.jp/cDNA/>  
NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.  
FAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niihara, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carrinchi, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, E., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.  
Location/Qualifiers  
1. 2343  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J013059F15"

Query Match 24.2%; Score 669.6; DB 8; Length 2343;  
Best Local Similarity 66.8%; Pred. No. 2.8e-66;  
Matches 1195; Conservative 0; Mismatches 424; Indels 171; Gaps 10;  
ORIGIN

165 GCGGCCATGATACCGCGTGGACCTACACGCTGCTACCGCGGTGCGGTGTAC 224  
145 GCGAAGATGATTACGCGCGGACTTCTACAGTATGACGCGGATGTCGTTGTAC 204  
225 GTGGCCATGACGCTGCGTACGCTCCGTCGCTGCTGCGGATCTTACGCGGACCG 284  
205 GTGGCGATGATACGCGTACGCTCCGTCGCTGCTGCGGATCTTACGCGGACCG 264  
285 TGCTCCGGATCAACCGCTCGTGGGCTCTGCGGCTGCGGCTCTCTCTCCACTTC 344  
265 TGCTCCGGATCAACCGCTCGTGGGCTCTTCCGCGTGGGCTCTGCTGCTCTCTC 324  
345 ATCTCCACCAACGACCCCTTCGCGCATGAACCTGCTCTCTGCGCGCGGACGCTG 404  
325 ATCTCCACCAACGACCCGCTACGATGACCTCGGTTTCATGCGCGCGGACGCTG 384  
405 AAGTGGCCCTCTCGCGCTGCTGGGCTGGGCTCCCGGCGCTCTCTCTCCCGCGCG 464

1437 ATCCCTCATGCTGGTGGAGGAGCTGATCCGGACCCACACCTACTCTCAGGCTCATC 1496  
1462 ATCCCTGATCATGCTGGTGGGAGAGCTATCCGACACCACTACTCTCAGGCTCATC 1521  
1497 GCGGTGCTGCTGGTCCCTGGTCTTCTACAGGTGGGCGCATCGAGATCGCATCGCC 1556  
1522 GCGCTCATGCTGGTCCCTGGTCTTCTACAGGTGGGCGCATCGAGATCGCGGCTCCTG 1581  
1557 CGGTGATTTTCATCTCTCGCGAGCGGGTCTCGGGATGGCCATGTTTCAGCTAGGCGCTG 1616  
1582 AAATCCATCTGATCTCTCGAGCGGGGCTCGGCATGGCCATGTTTCAGTCTCGTCTG 1641  
1617 TTCATGGCGCTCGACCGCAGGATCATCGCTGCGGGACACAGCTGGGGCGCATCGGATG 1676  
1642 TTCATGGCGCTCGACCGCAGCAGCATCATCGCTGCGGGACACAGTGGCGCATCGCCATG 1701  
1677 GCGTCCGGTTCGTCGCGAGCGCCCGCTGATGCGCGCGCTCCATCGCGCTGGTCTG 1736  
1702 GCGGTGCGGTTCGTCGCGCGCGCGCTGATGCGCGCGCTGCTTCGCGCTCGGACTC 1761  
1737 CGCGCGCTCTCTCCATCGCATCGCTGCGAGGCTGCTGCTCGCTCAGGGATCGTCCCG 1796  
1762 CGTGGCAGCTCTGTCAGCTGCGCATTTGTCAGGCGCTGTCGCCAGGCGATTGTCGCC 1821  
1797 TTCGTGTCGCGAGGATAGGCTGTCATCCGACATCTCTGACACAGC 1846  
1822 TTCGTCTTCGCGTAGGATCAGCGTGCACCTAGCATTTCTCAGCAGCAGC 1871

RESULT 11  
LOCUS AK101504 2457 bp mRNA linear PLN 24-JUN-2003  
DEFINITION Oryza sativa (japonica cultivar-group) cDNA clone:J033044E23, full insert sequence.

ACCESSION AK101504  
VERSION AK101504.1 GI:32986713  
KEYWORDS FLI CDNA; CAP trapper.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Eriarthroideae; Oryzaceae; Oryza.

1 The Rice Full-Length cDNA Consortium, National Institute of Agrobiological Sciences Rice Full-Length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, K., Kurokawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohseuki, K., Shishiki, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group, Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M., Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Niikura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN, Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Ootomo, Y., Sato, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Yoshino, M., and Hayashizaki, Y.  
japonica rice  
Collection, mapping, and annotation of over 28,000 cDNA clones from Science 301 (5631), 376-379 (2003)

JOURNAL japonica rice  
MEDLINE Science 301 (5631), 376-379 (2003)  
PUBMED 22752273  
12869764

REFERENCE 2 (bases 1 to 2457)  
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Hotta, I., Iida, J., Iida, J., Ikeda, R., Imamura, K., Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katch, H., Kawagashira, N., Kawai, J., Kawamata, M.,

Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurokawa, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K., Ooka, H., Ootomo, Y., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, K., Sakai, K., Shinagawa, A., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shizagawa, N., Shira, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyata, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.  
Direct Submission  
Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agrobiological Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (e-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
This clone is one of the 28K full-length cDNA clones from japonica rice.

URL : http://cdna01.dna.affrc.go.jp/cDNA/  
NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohseuki, K., Shishiki, T. and Yamamoto, M.  
PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katch, H., Kawai, J., Kouda, M., Kishikawa-Hirozane, T., Kojima, K., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ooka, H., Ootomo, Y., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers  
1..2457  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033044E23"

FEATURES  
source  
1..2457  
Query Match 22.8%; Score 632.2; DB 8; Length 2457;  
Best Local Similarity 63.2%; Pred. No. 4.5e-62;  
Matches 1141; Conservative 0; Mismatches 538; Indels 126; Gaps 6;

ORIGIN  
Query 163 CCCTGGCCATGATCATCGGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 222  
Best 344 CTGAGGAGATGATATCCGGGCGACGACTTCTACACGGTGATCGCGGGTGGCGGCTGT 403  
Query 223 ACCTGGCCATGATCATCGGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 282  
Best 404 ACCTGGCGATGTTCTTGGCGTACGGGTGGTGGCGGCATCTTCACGCCGGACC 463  
Query 283 AGTGTCTCCGGATCAACCGCTTCGTGGGCGCTCTTCGCCGTCGCGGCTCTCTCTCCACT 342  
Best 464 AGTGTCTCCGGATCAACCGCTTCGTGGGCGCTCTTCGCCGTCGCGGCTCTCTCTCCACT 523

FEATURES  
source  
1..2457  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J033044E23"

ORIGIN  
Query Match 22.8%; Score 632.2; DB 8; Length 2457;  
Best Local Similarity 63.2%; Pred. No. 4.5e-62;  
Matches 1141; Conservative 0; Mismatches 538; Indels 126; Gaps 6;

ORIGIN  
Query 163 CCCTGGCCATGATCATCGGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 222  
Best 344 CTGAGGAGATGATATCCGGGCGACGACTTCTACACGGTGATCGCGGGTGGCGGCTGT 403  
Query 223 ACCTGGCCATGATCATCGGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 282  
Best 404 ACCTGGCGATGTTCTTGGCGTACGGGTGGTGGCGGCATCTTCACGCCGGACC 463  
Query 283 AGTGTCTCCGGATCAACCGCTTCGTGGGCGCTCTTCGCCGTCGCGGCTCTCTCTCCACT 342  
Best 464 AGTGTCTCCGGATCAACCGCTTCGTGGGCGCTCTTCGCCGTCGCGGCTCTCTCTCCACT 523

FEATURES  
source  
1..2457  
Query Match 22.8%; Score 632.2; DB 8; Length 2457;  
Best Local Similarity 63.2%; Pred. No. 4.5e-62;  
Matches 1141; Conservative 0; Mismatches 538; Indels 126; Gaps 6;

ORIGIN  
Query 163 CCCTGGCCATGATCATCGGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 222  
Best 344 CTGAGGAGATGATATCCGGGCGACGACTTCTACACGGTGATCGCGGGTGGCGGCTGT 403  
Query 223 ACCTGGCCATGATCATCGGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 282  
Best 404 ACCTGGCGATGTTCTTGGCGTACGGGTGGTGGCGGCATCTTCACGCCGGACC 463  
Query 283 AGTGTCTCCGGATCAACCGCTTCGTGGGCGCTCTTCGCCGTCGCGGCTCTCTCTCCACT 342  
Best 464 AGTGTCTCCGGATCAACCGCTTCGTGGGCGCTCTTCGCCGTCGCGGCTCTCTCTCCACT 523

FEATURES  
source  
1..2457  
Query Match 22.8%; Score 632.2; DB 8; Length 2457;  
Best Local Similarity 63.2%; Pred. No. 4.5e-62;  
Matches 1141; Conservative 0; Mismatches 538; Indels 126; Gaps 6;

ORIGIN  
Query 163 CCCTGGCCATGATCATCGGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 222  
Best 344 CTGAGGAGATGATATCCGGGCGACGACTTCTACACGGTGATCGCGGGTGGCGGCTGT 403  
Query 223 ACCTGGCCATGATCATCGGCTGGGCTACGGCTCCGCTGGTGGCGGCATCTTCACGCCGGACC 282  
Best 404 ACCTGGCGATGTTCTTGGCGTACGGGTGGTGGCGGCATCTTCACGCCGGACC 463  
Query 283 AGTGTCTCCGGATCAACCGCTTCGTGGGCGCTCTTCGCCGTCGCGGCTCTCTCTCCACT 342  
Best 464 AGTGTCTCCGGATCAACCGCTTCGTGGGCGCTCTTCGCCGTCGCGGCTCTCTCTCCACT 523

343	QY	TCATCTCCACCAAGCACCCCTTCGCCATGAACCTGGCTTCTTGGCGCGCGACACGCTGC	402
524	Db	TCATCTCCACCAAGCACCCGTAAGCCATGAACCTCGCTTCCTGGCGCGGACACGCTGC	583
403	QY	AGAAGGTGGCCGCTCTCGCGCTGTGGCGCTGGCGCTCCCGCGGCTCTCGTCCCGCGCG	462
584	Db	AGAAGCTGTCTGCTCTGGCGGGCTCGCGCGCTGGTGGCGGCTCCCTTCGCGACCGGG	643
463	QY	CGTTCGGGCTCGACTGGAGCATCACGCTCTTCTCCCTCTCCACGCTCCCCAAACACGCTCG	522
644	Db	CGCGCGCTGGACTGCTCCATCACGCTCTTCTCCCTCTCCACGCTGCCCAACACGCTCG	703
523	QY	TCATGGGCATCCGCTGCTCGAGGCAATGACGGCGGTGTCGGCGGGACGCTCATGG	582
704	Db	TCATGGGGATCCCGCTCTGATGCCATGTACGG---GCCATACTTCGGCTCGCTCATGG	760
583	QY	TCAGGTCTGCTCTCCAGTGCATCATCTGGTACACGCTCATGCTCTTCTCTTTCGAGT	642
761	Db	TCAGATCTGCTGCTTCAGTGCATCATCTGGTACAGCTGATGCTCTTCCACTTCGAGT	820
643	QY	ACCGCGCGCGCGCGCTCTGCTCTGACACAGTTCCCGACGGCGCGCGCGCTCCATCG	702
821	Db	TCGCGCGCGCGGATGCTGATCGCCACAGTTTCGG---CACGGCGGCGTCCATCG	877
703	QY	TCCTCTTCGGCTCGACTCCGAGTCTCTGCTGCTCGCACGGGGGAGCTCGASCTCGAGG	762
878	Db	TGTCCTCTGCACTCGACCCGACGTGTGTCTGAGGGCGCGCACCGGAGACGAGG	937
763	QY	CCGAGCCCGAGCGGCTGCGCGCGCCCGCGCGCTCTCTCCGCGCGGGAGCGCCGGGC	822
938	Db	CCGAGGTGGCGGAGACGGCGGGCTGACGCTCACGCTGCGCGCGCTCTCGGTGTCGCGC	997
823	QY	GGGTGCGGCTCACGTCGGCAAGTCCACAGCTCGGCTCGGAGGCGCGTGTCTGCACCT	882
998	Db	GGTCGCTGCTGTTCACGCGCGGCGCTGCAACCTGACGGAGCGGAGATCTATCTCGCTTA	1057
883	QY	CGCACTCCAGAACCATGAGCCCGCTGTGTCCAACTCTCCGCGTGGAGATCTATCTCGC	942
1058	Db	GCTCGTTCGGGAACCAACCCGCGGGCTCCAACTTCAACACGCGCATCTTCTCGCCA	1117
943	QY	TGCAGTCTGCGCAACCCCAACCCGCGGGGTCCAGTTCACACCGCGCATCTTCTTCA	1002
1118	Db	TGGTCGGCGGGGCCACCGCCCGACGCGCGCTGCGGTGCGGGCTCGAGCTTCGGCG	1177
1003	QY	ACATCGTCGGCGCGCGCCCAAGGAGGCGGAGAGCGGGGGACGAGGAGAGGGCG	1062
1178	Db	CCTCGAGCTCTATCTCGTGCAATGCTGCGGGGCCCAACCCGAGGCACTTCCATCTCG	1237
1063	QY	CATGCGCGGGCGGGGAGACACTCGCGCAGCCGCGAGCGCGTGCCTGCGCGCCA	1122
1238	Db	ACGAGCACTCGGCACGGCCACCGAAACACCGGCAACGACCGGGGCACTCAACACCG	1297
1123	QY	AGAGGAAGGACTGCATGCTCTGTGGAGCTCCAGGGCTGCGCCGCTGTCGAGCGCG	1182
1298	Db	ATGCCAAGGAGCTCCATGTTCTGTGGAGCTGAGCGGCTATCCCGTCTTCAGAGTCA	1357
1183	QY	CGCGCTGCACTCTTC---GGCGCGCGCGCGCTGACCATGCGCAGAGTCTCTCGCCAAAG	1239
1358	Db	CGCGCTGCTGTGTTCAGTGGCGGCGCGCGCGGCTCTCGAGCTCGGCGCCAGG	1417
1240	QY	GAGGCCA-----	1246
1418	Db	AAATCCACATGTCATCCCGCGACCTTCGCGCAGAACACGCTCAGGCAAGAGCACG	1477
1247	QY	-----GGCTACGACGATACGGGCGCGACGACTACA	1278
1478	Db	AGGAGTACGGCGAGTGGCATTGGTGGCGGGCGCGGAGAGAACTTCAGCTTCGGAG	1537
1279	QY	GCAGCAGGACGAAGAACGGGAGCGCGCGCGGACGAGGGCGGGCC-----GACGC	1329
1538	Db	CGCGCAAGACGCTGTGACCGCGCGGAGGAGTAGACGAGGAGCGGCGCTTGCTGACGGCG	1597

japonica rice  
 Science 301 (5631), 376-379 (2003)  
 22752273  
 12869764  
 2 (bases 1 to 2276)  
 ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
 Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiroaka, T.,  
 Hori, F., Hotta, I., Ikeda, Y., Ikeda, Y., Ikeda, K., Imamura, K.,  
 Imotani, K., Ishibiki, J., Ishii, Y., Itoh, M., Itoh, M., Kagawa, I.,  
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawata, M.,  
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
 Mizuno, K., Murakami, K., Murata, M., Negata, T., Nakamura, M.,  
 Nami, T., Nariakawa, R., Niikura, J., Nishi, K., Nomura, K.,  
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
 Osato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
 Shinagawa, A., Shiraki, T., Shisiki, T., Sogabe, Y., Sugano, S.,  
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
 Toya, I., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanaga, W.,  
 Yanada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and  
 Yoshimura, A.  
 Direct Submission  
 Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of  
 Agrobiological Sciences, Department of Molecular Genetics, Head of  
 Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki  
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,  
 Tel:81-29-838-7007, Fax:81-29-838-7007)  
 This clone is one of the 28K full-length cDNA clones from japonica  
 rice.  
 URL : <http://cdna01.dna.affrc.go.jp/cDNA/>  
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,  
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
 Ishikawa, M., Yanada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., and  
 Ohneda, E., Yanaga, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and  
 Yamamoto, M.  
 FAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y.,  
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawata, M., Kobayashi, M.,  
 Kodama, T., Kurotaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
 Mizuno, K., Nariakawa, R., Niikura, J., Oka, M., Ryu, R., Sugano, S.,  
 Sugiyama, A., Suzuki, Y., Tanoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
 Yoshimura, A., Matsubara, K. and Murakami, K.  
 Genome Exploration Research Group in Riken Genomic Sciences Center  
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
 Hiroaka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,  
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
 Yasunishi, A. and Hayashizaki, Y.  
 Location/Qualifiers  
 1..2276  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultiVar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J033030G03"

Db	1118	TCCAACTCCAAACAGGAGCTACACATGTTGCTGTGGAGCTCTAGCGCGTGC	1177
QY	1176	GAGCGCGCCGCGTGACGCTCTTCGGCGCGCGC-----	1208
Db	1178	GAGGCCAACTCCGCAACGCGGTCAACACGCGCGCTCCACCGACTTCGCGCTCGCGCG	1237
QY	1209	-----GGCGCTGACCAATGCCGAGCTCTCTGCGCAAAGAGGCCCGAGGCTTACGACGAG	1259
Db	1238	CCGCGCGACCGGTTCCTCGTCCGCGCGCGCACTCCCAAAGGGGTGAGTGGCACTGTCA	1297
QY	1260	TACGGGCGCACGACTACAGCAGCAGACGAAGAACGGGAGCGCGCGCGGACAAAGGCG	1319
Db	1298	CCGCGCGGAAGAACGCGCGCGCGGAGTTGAGATCGAGACGCGGCTGAAGAGCCCGCG	1357
QY	1320	GGGCGGACGCTGCGAAGCTGGGGTCCAACTCGACGCGCGAGTGTATCCCAAGGACGAC	1379
Db	1358	GCGGGGGTGCGCGCGAAGTTCCCGGTCTCGGGGTTCGCGGTACGTGGCGCGAGAGAAAG	1417
QY	1380	GGCGAGGG-----CAGGCGCGCGCGGTGGCGATGCCCGCGCGGAGC	1421
Db	1418	GGCGCGCGCGCGACGTGCCCGGCTGGCGAGCGCGCGCACCGATGCCCGCGACGAC	1477
QY	1422	GTGATGACGCGGCTCATCTCTCATCATGGTGTGAGGAAGTGATCGGGAACCCCAACACC	1481
Db	1478	GTGATGACGCGGCTCATCTCTCATCATGGTGTGCGCAAGCTCATCAAAAACCCCAACACC	1537
QY	1482	TACTCCAGCCTCATCGCGCTGCTGCTGCTGCTCTCTTACAGGTGGGSCATCGAGATG	1541
Db	1538	TACTCCAGCCTCATCGCGCTGCTGCTGCTGCTCTCTTCAAGTGGAAATATCCAAATG	1597
QY	1542	CCAGCGATATCGCGCGGTGATTTGCATCTCTCGACGCGGTCTCGGATGGGCATG	1601
Db	1598	CCTTCAATAAAGGGCTCAATATCAATATTGTTCAGATCAGGGCTAGGAATGGGTATG	1657
QY	1602	TTCAGCCTAGCCCTGTTTCATGCGGCTCGACCGAGGATCATCGGTGCGGGACAGCTG	1661
Db	1658	TTCAGCTTAGGCTTGTTTCATGGTCTCGAACCAAGATCATTTCTTGTGGCAAGCCGTT	1717
QY	1662	GCGCCATCGGATGGGCGTCCGGTTCGTGCGAGCGCCCGCGGTCTATGGCGCGCGCTCC	1721
Db	1718	CGACATTTGCAATGGCAGTGAGGTTCTTGACTGGTCCAGCTGTATTGCACTACTTCC	1777
QY	1722	ATCGCGGTGGTGTGCGCGGGTCTCTCTCGACATCGCCATCGTCCAGGCTGCTCTGCT	1781
Db	1778	ATTGCCATTGGGCTCAGGGAGTACTCTTGCTGATTTGCCATTGTTTCAGGAGCACTTCCA	1837
QY	1782	CAGGGGATCGTCCGTTGTTTTCGCAAGAGTACGGCGTTCACTCCGACATCTCTGAGC	1841
Db	1838	CAAGCGATTGTCCGTTTGTGTTTGCAGAGGATCAATTGGCCATCTCTCAATACTTAGC	1897
QY	1842	ACAAGC 1847	
Db	1898	ACAGCG 1903	

RESULT	13
AX046850	
LOCUS	AX046850 linear DNA 2293 bp PAT 15-DEC-2000
DEFINITION	Sequence 37 from Patent WO0068389.
ACCESSION	AX046850
VERSION	AX046850.1 GI:11876332
KEYWORDS	Triticum aestivum (bread wheat)
SOURCE	Triticum aestivum
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
REFERENCE	Poideae; Triticeae; Triticum. Crocco,E.M., Weng,Z., Bruce,W.B., Cahoon,R.E. and Tao,Y. Auxin transport proteins Patent: WO 0068389-A 37 16-NOV-2000; E.I. DU PONT DE NEMOURS AND COMPANY (US) ; Pioneer Hi-Bred
JOURNAL	

```

Db 949 GGGCCAGCCCCCGTCCGCGAGGTGGTGGCGCGGCGAGGGCTCGACGAGCAGGT 1008
QY 1058 GGGCGCATCGCGCGCGCGGCGAGAGACTCCCGCAGCGCGAGCCGTCGCCGTG-- 1115
Db 1009 GGCCAAACAAGTTCAAGGCGCGGCGAGGCGGTGCGCCCTTACCCCGCCCAACCCCGGAT 1068
QY 1116 -----CCGCGCCCAAGAGGAGGA 1132
Db 1069 GATGATCCCGCGCCACCGAAGAGGAGCTTGGGGTTCCAACTCAACTCGACRAGGA 1128
QY 1133 CTTGCACATGCTGTTGGAGCTCAGCGCTCCCGCGTGTCCAGCGCGCGCGCTGCA 1192
Db 1129 CTTGCACATGTTCTGTGGAGCTCAGCGCTCCCGGTGTGCGAGGCGCAACTCCGCAA 1188
QY 1193 ---CGTCTTCGCGCGCGCGGTGACCATGCGGAGTCTCTGCCAAAGAGCCAGGC 1249
Db 1189 CGCGCTCAACCAACCGCGGTCCACGACTTCGCGCGCGCACCGCGCGGAGCCAGCC 1248
QY 1250 CTACGACAGTACCGGCGCGACGACTACAGCAGCAG-----GACGAAGAA 1294
Db 1249 ACGAGACGGGCCACACCCAGAGCGTGGAGCGGAGCGTGACGCGGTGATGAAGAGGA 1308
QY 1295 CGGAGCGCGCGCGCGGACAGAGGG-----GGCGAGCGTGC 1333
Db 1309 CGCAGCAGCGCGCGGTGGAGGTGGAGATCGAGAGCGGCGATGAGAGAGCCGCGCAG 1368
QY 1334 GAAGCTGGGGTCCAACTCGACGGGCGAGCTGTACCCCAA----- 1372
Db 1369 GGGCTGGGCGCGCAAGTTCCCGGTGTCGGGTCCCGCTACGTGGCCCGCGGAAGAGG 1428
QY 1373 ---GGAACAACCGGAGGAGGCGCGCGGTGGCGATCGCGCGCGGCGAGCGTATGAC 1429
Db 1429 CGCCACATGCTGCGGTGAGGAGGCGCGCACCGATGCGCGCGGCGAGCGTATGAC 1488
QY 1430 GCGCTCATCTCATGATGATGAGGAGAGCTGATCGGAACCCCAACACTACTCCAG 1489
Db 1489 CCGGCTCATCTCATGATGAGGAGAGCTGATCGGAACCCCAACACTACTCCAG 1548
QY 1490 CTTATCGCGGTGCTGCTGCTCCCTGCTCTACAGGTGGGGCATGAGATGCGAGCGAT 1549
Db 1549 CTTATCGCGGTGCTGCTGCTCCTGCTCCTTACAGTGGAACTTACAGATGCTCAAT 1608
QY 1550 CATCGCGCGGTGATTTGATCTCTGTCGAGCGGCTCTCGGATGCGCATGTTACGCT 1609
Db 1609 AATCAAGGGGTCCATATCCATCTCTGTAAGAGGCTAGGATGCTATGTTACGCTT 1668
QY 1610 AGGCTGTTTCATGGCGCTGACCGGAGGATCATCGCTGCGGGAACAAGCTGGCGGCAT 1669
Db 1669 AGGTCTCTTCATGGCTCTGCAACCAAGATCATCTTTGCGGGAAGTCTGCGCCACAT 1728
QY 1670 CCGATGGGCGTCCGCTGCTGCGAGGCGCGCGGTGATGCGCGCGCTTCACTCGCGT 1729
Db 1729 TGCAATGGCAGTGAAGTCTTTGACTGGCGCGGTGATGCGCGCGCTCAATCGCGGT 1788
QY 1730 CCGTCTGCGCGGTCTCTCTCCATCGCATCGCTGCGAGGCTGCTCTGCTCAGGGGAT 1789
Db 1789 CCGGCTCCGGGAGTCTCTCTACATGTTGCCATTGTCAGGAGCACTTCCACAGGAT 1848
QY 1790 CTTGCGGTGCTGTTGCGCAAGAGTACGCGGTTTCATCCCGACATCTGAGCAGCAGG 1847
Db 1849 TGTTCATTTGTGTTGCGCAAGAGTACAATTTGCCATCTCTCAATACTTAGCAGCG 1906

```

## RESULT 14

BT008949

LOCUS

DEFINITION

BT008949

ACCESSION

VERSION

KEYWORDS

SOURCE

2293 bp mRNA linear PLN 20-JUN-2003  
Triticum aestivum clone wdk1c.pk008.g12.fis, full insert mRNA

sequence.

BT008949

BT008949.1

FLI\_CDNA.

Triticum aestivum (bread wheat)

## ORGANISM

Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Poideae; Triticeae; Triticum.

## REFERENCE

AUTHORS

Tingey,S.V., Wolters,P., Powell,W., Dolan,M., Miao,G.-H.,

Caraher,N.R., Hanafey,M.K. and Hainey,C.F.

## TITLE

Direct Submission

Submitted (20-JUN-2003) Crop Genetics, E. I. DuPont de Nemours and

Company, 1 Innovation Way, P.O. Box 6104, Newark, DE 19714-6104,

USA

## FEATURES

source

/organism="Triticum aestivum"

/mol\_type="mRNA"

/db\_xref="taxon:4565"

/clone="wdk1c.pk008.g12.fis"

## ORIGIN

Query'Match 22.6%; Score 626.8; DB 8; Length 2293;  
Best Local Similarity 64.5%; Pred.No.1.8e-61;  
Matches 1199; Conservative 0; Mismatches 467; Indels 192; Gaps 10;

QY 164 CCGCGCATGATCACCGCGCTGACCTCTPACCACTGTGTGAGCGGCTGGTCCGCTGTA 223  
Db 67 GTCTGGGATGATACCGGAAGACATCTACGACGTGTGGCGGCGGTGGTCCGCTGTA 126  
QY 224 CTTGCGCATAGCCTGGGTACGGCTCCGTCCGTGGTGGCGCATCTTCACCGGACCA 283  
Db 127 CTTGCGCATGTTTCATGGCGTACGGGTCCGTGGGTGGTGGGCACTTTCACCGGACCA 186  
QY 284 GTGTCCGGGATCAACCGCTTCTGGCGCTCTTCGCGTCCGCTCCCTCTCCACTT 343  
Db 187 GTGCTCGGCGATCAACCGCTTCTGGCGCTCTTCGCGTCCGCTCTCTCTCCACTT 246  
QY 344 CATCTCCAAACGACCCCTTCGCCATGAACCTGGCTTCTTGGCGCGCACACGCTGA 403  
Db 247 CATCTCCAAACGACCCCTTCGCCATGAACCTGGCTTCTTGGCGCGCACACGCTGA 306  
QY 404 GAAGTGGCGCTCTCGCGCTGCTGGCGTGGCTCCCGGCGCTCTCTCCCGCGCGC 463  
Db 307 GAAGTGGCGCTCTCGCGCTGCTGGCGTGGCTCCCGGCGCTCTCTCCCGCGCGC 366  
QY 464 GTCGGG-----GTCGACTGGAGCATCACGCTCTTCTC 496  
Db 367 CCGCGCGGACCGAGGCGCGGAGGCTCTGCTGGACTGGACCATCACGCTCTTCTC 426  
QY 497 CTTCTCCAGCTCCCAACACAGCTCTGTATGGGATCCCGCTGCTGCGAGGCTATACGG 556  
Db 427 CTTGCGAGCTGCCCAACACAGCTGTGTATGGGCTCCCGCTGCTGCGGCGCATGACGG 486  
QY 557 CCGGCTCTCGCGCGCACGCTCATGTGCCAGTCTGTCCTCCAGTGCATCATCTGGTA 616  
Db 487 GCACTTCTCG---GGTCTCATGTGAGATCTGTGTGCTGAGAGGCTCATCTGGTA 543  
QY 617 CAGCTCATGCTTTCTCTTTTGAAGTACCGCGCGCGCGCTGCTCTTCTGACAGTT 676  
Db 544 CAGCTCATGCTTTCTCTTTTGAAGTACCGCGCGCGCGCTCATCTGAGAGT 603  
QY 677 CCGCGAGGCGCGCGCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 736  
Db 604 CCGCGCGAGCTGCGCGCGCAGCATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 663  
QY 737 CCGCAGGCGGAGCTGAGCTGAGGCGGAGCGGCGGCTGCGCGCGCGCGCGCGCT 796  
Db 664 CAACGGCGCGAGGCGCTGACGCGGAGGTCGGCGCGGCGGCGGCGGCTCCACGT 723  
QY 797 CTCTCTCGCGCGGAGGAGCGCGGCGGCTGCGGCTCACGCTGCGCAAGTCCACAGCTC 856  
Db 724 CGTCTCTCGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 783  
QY 857 GCGCTCGGAGGCGCGGCTGCTCGCACTCGCACTCCAGACCATGACGCTGCTGCTCAA 916

Db	784	GATCTACGGTGGCGG-----TCCAAAGCCCATGACGCCGGCGCGTCCAA	828
Qy	917	CCTCTCGGCGTGGAGATCTACTCGTCTGAGTCTGTCGCGCAACCCCAACCCCGCGCGGGTCC	976
Db	829	CCTCACGGCGTGGAGATCTACTCGTCTGAGACGTCTGCGGAGCACCACGCCGAGGCGAGTC	888
Qy	977	CAGCTTCAAACAACGCCGACTTCTTCAACATCGTCGGG-----1014	1014
Db	889	CAGCTTCAACCATGCTCCGACTTCTTACTTCATGTTTCAAACGGGAGCAAGCTGGCTAGTCCCAA	948
Qy	1015	-----CGCCGCCCAAGGAGCGGAGGACGCGCGCGG-----GACGAGAGAA	1057
Db	949	GGGCCAGCCCCCTCGTCCGAGGTGTGTGTGCGCGCGGACAGGGCTCTGACGAGCAGGT	1008
Qy	1058	GGGCGCATGCGGCGGCGCGCGGAGAGACTCTGCGCGAGCGCGAGAGCCGTGCCGTG--	1115
Db	1009	GGCCAAACAAGTTCAGGGCGGCGAGGCGGTCTGCGCCCTACCCCGCGCCAAACCCCGGAT	1068
Qy	1116	-----CGGCGCCAAAGAGGAAGGA	1132
Db	1069	GATGATGCCGGCGCCACGGAGAAGGAGCTTGGGGTTCNACTCAACTCGAACAGGA	1128
Qy	1133	CCTGCAATGCTCGTGTGAGCTCCAGCGCTCTGCCCGTGTCTGAGCGCGCCGCGTGC	1192
Db	1129	GCTGCATGTTCTGTTGTGAGCTCCAGCGCTCTGCCCGTGTCTGAGAGCCAACTCCGCAA	1188
Qy	1193	---CGTCTTTCGGCGCGCGCGCTGACCAATCCGAGAGTCTCTGCCAAAGAGGCCAGGC	1249
Db	1189	CGCCGTCAACCAACGCGCGGTCCACCGACTTCGCGCGCGCACCGCGCGGCGAGCACGCC	1248
Qy	1250	CTACGACGAGTACGGGGCGGACGACTACAGCAGCAG-----GACGAAGAA	1294
Db	1249	ACGAGACGGGCCACACCCAGAGCGTGGCGGAGCGTGACGCGGTGATGAAGAAGGA	1308
Qy	1295	CGGAGCGGCGCGCGGACAAAGGG-----GGCGAGCGTGTCT	1333
Db	1309	CGCCAGCAGCGCGCGGTGGAGGTGGAGATCGAGGACGGCATGATGAAGACCCGCGAC	1368
Qy	1334	GAAGCTGGGGTCCAACTCGACGGCGCGAGCTGTACCCCAA-----1372	1372
Db	1369	GGGCTGGCGCCAAAGTTCCTCGGTCTGGGTCTCCCTTACGTGCCCGCGGAGAGGG	1428
Qy	1373	---GGAAGAAGCGGAGAGGAGGGCGGCGCGGTGGCGATGCCCGCGGAGCGTATGAC	1429
Db	1429	CGCCGACGTGCTCGGGCTGGAGAGGCGGCGCACCGATGCGCGCGGCGAGCGTATGAC	1488
Qy	1430	CGGCTCATCTCATATGTTGTGAGGAAGCTCATCCGGAACCCCAACACTTACTCCAG	1489
Db	1489	CGGCTCATCTCATATGTTGTGGCGAGCTCATCCGCAACCCCAACACTTACTCCAG	1548
Qy	1490	CCTCATCGCGCTCGTCTGGTCCCTTGTCTCTTACAGTGGGGCATCGAGATGCCAGGAT	1549
Db	1549	CCTCATCGCGCTCGTCTGGTCACTCGTCTCTTACAGTGGGAACATTGATGCCATTAAT	1608
Qy	1550	CATCGCCCGTTCGATTCGATTCCTTCGGAAGCGGGTCTCGGATGGCATGTTTCAGCT	1609
Db	1609	AATCAAGGGGTCCATATCCATCTCTGTATGCAAGGCTGAGGATGGCTATGTTTCAGCT	1668
Qy	1610	AGGCTGTTTCATGGCGCTGACGCGAGGATCATCGCTGCGGGGAACAGCTGGCGGCAT	1669
Db	1669	AGGTCTCTTCATGGCTCTGCAACCAAGATCATCTTTCGCGGAAGTCTGTGCGCAATT	1728
Qy	1670	CGGATGCGGCTCGGTTCTGTCGACAGGCCCGCGGTTCATGCGCGCGGCTCCATCGCGT	1729
Db	1729	TGCAATGACGTGAGGTTCTTGACTTGGCGCGCGGTGATCCCGCGACCTCAATTCGCGT	1788
Qy	1730	CGGCTCGCGCGGCTCTCTCCACATCGCCATCGTCCAGGCTGCTGTGCTCAGGGGAT	1789
Db	1789	CGGCTCGCGGAGTGCTCTCTACATGTTGCCATTGTCCAGGACGACTTCCACAGGAAT	1848
Qy	1790	CGTGCCTGCTGTTCGCCAAGGAGTACGGCGTTTCATCCCGACATCTCTAGCACAGCG	1847
Db	1849	TGTTCCATTGTTTTCGCAAGGAGTACAAATGGCATCTCAATGCTCAAAATCTATGACACAGG	1906

RESULT 15	AX654121	1827 bp	DNA	linear	PAT 22-MAR-2003
LOCUS	AX654121				
DEFINITION	Sequence 3991 from Patent WO03000898.				
ACCESSION	AX654121				
VERSION	AX654121.1	GI:29156935			
KEYWORDS					
SOURCE	Oryza sativa				
ORGANISM	Oryza sativa				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzoideae; Oryza.				
AUTHORS	Chang, H.S., Chen, W., Cooper, B., Glazebrook, J., Goff, S.A., Hou, Y.M., Kataqiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T. and Zou, G.				
TITLE	Plant genes involved in defense against pathogens				
JOURNAL	Patent: WO 03000898-A 3991 03-JAN-2003;				
FEATURES	Syngenta Participations AG (CH)				
source	Location/Qualifiers				
	1. 1827				
	/organism="Oryza sativa"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4530"				
ORIGIN					
	Query Match	22.3%	Score 618;	DB 6;	Length 1827;
	Best Local Similarity	63.2%	Pred. No. 1.9e-60;		
	Matches 1116;	Conservative	2;	Mismatches 522;	Indels 126; Gaps 6;
QY	202	TGACGGCGGTGGTGC	CGCGCTGTACGTGGCCATACGCTGGCGTACGCGTCCGTCGGTGGT	261	
DB	2	TGGCGCGGTGGTGC	CGCGCTGTACGTGGCGATGTTCTGGCGTACGCGTGGTGGTGGT	61	
QY	262	GGCGCATCTTCAGCGCGGACGAGTCTCCGGATCAACCGCTTCGTGGCGCTCTTCGCGG	321		
DB	62	GGGCGATCTTCACCGCGGACGAGTCTCCGGATCAACCGCTTCGTGGCGCTCTTCGCGG	121		
QY	322	TGCGCTCTCTCTCTTCCATTCATCTCCACCAACGACCCCTTCGCCATGAACCTCGCGCT	381		
DB	122	TGCGCTCTCTCTCTTCCATTCATCTCCACCAACGACCCGCTACGCCATGAACCTCGCGCT	181		
QY	382	TCCTGGCGCGGACGACGCTCAGAAAGTGGCGTCTCGCGCTGCTGGCGCTGCGCTGCCCTCC	441		
DB	182	TCTTGGCGCGGACACGCTGACGAAGTGTGCTCTTGGGGGGCTCGCGCGTGGTGGT	241		
QY	442	CGCGCTCTCTCTCCCGCGCGCGCTCGGGCTCGACTGGAGATCAGCGTCTTCTCCCTCT	501		
DB	242	GCTTCCCTCGCGACCGCGCGCGCGGCTGGACTGCTGCATCACGCTCTCTCTCCCTCT	301		
QY	502	CCAGCTCCCCAACACGCTCGTATGGGATCCCGCTGCTCGAGGATGTACGGCGGT	561		
DB	302	CCAGCTGCCAACACGCTGTCTATGGGATCCCGCTGCTGATCGCATGTACGG---	358		
QY	562	CGTGGCGCGGACGCTCATGGTCAAGTGTGCTGCTCCAGTGTGATCATCTGTGTACAGC	621		
DB	359	CATACCTCGGCTCGCTCATGGTCCAGATCGTGTGCTCCAGTGTGATCATCTGTGTACAGC	418		
QY	622	TCATGCTCTTCTCTTCAGTACCGCGCGCGCGGCTCGTCTCTCGACCAAGTTCGCGG	681		
DB	419	TGATGCTCTTCTCTTCAGTTCGCGCGCGCGGATGCTGATCGCGACCAAGTTCGCGG	478		
QY	682	ACGCGCGCGCGGTCTCATGCTCTCTTCGCGTGTGATCTCGAGCTGTCTCGCTCGCA	741		
DB	479	A---CAGGGGGGCTCCATCGTGTCTCTTCGACGCTCGACCCGCGAGTGTCTGTGGAGG	535		
QY	742	GGGGGAGCTCGAGCTCGAGCGCGAGCCGAGCGGCTCGCGCGCGCGCGCTCTCT	801		
DB	536	CGGCGCAGCGGAGACGAGAGCCGAGTGTGCGCGGACGGGGGCTGCACGTACCGTGC	595		
QY	802	CCC CGCGCGGAGCGCGCGCGGTGCGCTCAACGCTGCGCAAGTCCACAGCTCGCGCT	861		



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 3, 2004, 00:50:45 ; Search time 1036 Seconds  
(without alignments)  
11354.499 Million cell updates/sec

Title: US-10-030-884-13  
Perfect score: 2769  
Sequence: 1 ccacgcctcgctgagccct.....aaaaaaaaaaaaaaaaaaag 2769

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_29Jan04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2003as:\*  
8: geneseqn2003bs:\*  
9: geneseqn2003cs:\*  
10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2769	100.0	2769	3	AA94721 Corn auxi
2	838.6	30.3	1926	7	ADA70391 Rice gene
3	760	27.4	1686	7	ADA69470 Rice gene
4	683	24.7	1884	7	ADA70233 Rice gene
5	671.2	24.2	2374	3	AZ29280 Rice BIR1
6	626.8	22.6	2293	3	AA94733 Wheat aux
7	618	22.3	1827	7	ADA70668 Rice gene
8	563	20.3	1845	7	ADA69396 Rice gene
9	536.2	19.4	2162	3	AA94724 Rice auxi
10	452.6	16.3	504	3	AA94720 Corn auxi
11	425.6	15.4	1426	3	AA94719 Corn auxi
12	347.2	12.5	1851	3	AAC43329 Arabidops
13	329.8	11.9	2292	2	AAV16361 cDNA sequ
14	323.8	11.7	1088	3	AA94716 Corn auxi
15	319	11.5	1860	6	ABZ13043 Arabidops
16	319	11.5	1860	7	ADA68467 Arabidops
17	319	11.5	2027	3	AAC44892 Arabidops
18	316.6	11.4	2549	3	AA94729 Soybean a
19	305.6	11.0	1713	3	AAC48532 Arabidops
20	304.8	10.9	2101	3	AA94727 Soybean a
21	290.8	10.5	1923	3	AAC44722 Arabidops
22	290.8	10.5	1923	6	ABZ13042 Arabidops
23	267.6	9.7	1725	3	AAC46518 Arabidops

24	254.8	9.2	2204	3	AAZ29279	Aaz29279 A. thalia
25	254.2	9.2	2061	3	AAC46177	Aac46177 Arabidops
26	254.2	9.2	2232	3	AAZ57348	Aaz57348 Gravitrop
27	251	9.1	2117	2	AAV16362	Aav16362 cDNA sequ
28	249	9.0	1618	3	AA94725	AA94725 Soybean a
29	232.6	8.4	855	3	AA94737	AA94737 Corn auxi
30	232	8.4	1983	6	ABZ13500	Abz13500 Arabidops
31	226	8.2	2324	3	AA94731	AA94731 Soybean a
32	209	7.5	415	3	AA94735	AA94735 Wheat aux
33	209	7.5	425	3	AA94736	AA94736 Wheat aux
34	195.2	7.0	473	3	AA94732	AA94732 Wheat aux
35	195	7.0	629	3	AA94715	AA94715 Corn auxi
36	176.6	6.4	330	3	AA94723	AA94723 Rice auxi
37	168.8	6.1	624	3	AA94718	AA94718 Corn auxi
38	158.2	5.7	543	3	AA94722	AA94722 Rice auxi
39	148	5.3	531	3	AA94726	AA94726 Soybean a
40	140.2	5.1	447	3	AA94734	AA94734 Wheat aux
41	139	5.0	3980	3	AAZ57349	Aaz57349 Gravitrop
42	139	5.0	7072	3	AAZ29278	Aaz29278 A. thalia
43	130.6	4.7	514	3	AAC41293	Aac41293 Zea mays
44	118.2	4.3	620	7	ABZ73099	Abz73099 Rice leaf
45	115	4.2	14055	7	AAU61170	Aal61170 Actinosyn

ALIGNMENTS

RESULT 1  
AAA94721  
ID AAA94721 standard; DNA; 2769 BP.  
XX AC AAA94721;  
XX AC  
DT 02-FEB-2001 (first entry)  
XX  
DE Corn auxin transport protein clone p0119.cmtnl24r DNA sequence.  
XX  
KW Auxin transport protein; corn; root development; gene mapping;  
KW plant breeding; herbicide; ss.  
XX  
OS Zea mays.  
XX  
PN WO2000069389-A2.  
XX  
PD 16-NOV-2000.  
XX  
PF 03-MAY-2000; 2000WO-US012061.  
XX  
PR 07-MAY-1999; 99US-0133040P.  
XX  
PA (DUPO ) DU PONT DE NEMOURS & CO E I.  
PA (PION-) PIONEER HI-BRED INT INC.  
XX  
PI Orococo EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
XX  
WPI: 2000-687647/67.  
DR P-PSDB; AAB26934.  
XX  
XX New nucleic acid sequences encoding new auxin transport proteins, useful  
for modulating root growth of plants and to screen for herbicides.  
XX  
XX Claim 2; Page 60-61; 94pp; English.  
XX  
XX Auxins are plant hormones that influence plant behaviour and development  
e.g. vascular tissue differentiation, apical development, tropic  
responses and organ (e.g. flower, leaf) development. The present  
invention relates to corn auxin transport protein coding sequences and  
proteins. The present sequence is one such coding sequence. This sequence  
may be used to modulate root development, e.g. to produce a more robust  
root system, alter root angle or redirect root growth. Also, the present  
sequence may be useful for gene mapping (e.g. for plant breeding) and to  
identify loss of function mutants. The protein encoded by the present  
sequence may be useful for raising specific antibodies, for the detection

CC of auxin transport proteins and to design and/or identify specific  
CC inhibitors of auxin transport proteins, potentially useful as herbicides  
XX  
SQ Sequence 2769 BP; 527 A; 917 C; 757 G; 568 T; 0 U; 0 Other;

Query Match	100.0%;	Score 2769;	DB 3;	Length 2769;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2769:	Conservative	0;	Mismatches	0;
	Indels	0;	Gaps	0;

Qy	1	CCACGCGTCGGTGAGCCGCTCAACACCTCTCTCTCTTCATCTGTCGCACACTACCACTCA	60
Db	1	CCACGCGTCGGTGAGCCGCTCAACACCTCTCTCTCTTCATCTGTCGCACACTACCACTCA	60
Qy	61	TCCTCCGCGCATTTTACACCACTCCCTCTCTGTTGCAACCAACAATTTGGCACTGCTCG	120
Db	61	TCCTCCGCGCATTTTACACCACTCCCTCTCTGTTGCAACCAACAATTTGGCACTGCTCG	120
Qy	121	TGCGGACCCCTCTCTCCCTCCGCGGTCCTCCCGACAAGCCATCCGGCGCATGATCACCG	180
Db	121	TGCGGACCCCTCTCTCCCTCCGCGGTCCTCCCGACAAGCCATCCGGCGCATGATCACCG	180
Qy	181	CGGTGACCTCTACACAGCTGTGACGGCGGTGGCGCTGTACGTGGCCATGACGCTGG	240
Db	181	CGGTGACCTCTACACAGCTGTGACGGCGGTGGCGCTGTACGTGGCCATGACGCTGG	240
Qy	241	CGTACGGCTCCGTCGCGTGGTGCGCATCTTTACGCCGGAACAGTGCTCGGGATCAAC	300
Db	241	CGTACGGCTCCGTCGCGTGGTGCGCATCTTTACGCCGGAACAGTGCTCGGGATCAAC	300
Qy	301	GCTTCGTGGCGCTCTTCGCGGTGCGCTCTCTCTTCCTCTTCCTCTTCCTCTTCCTCTTC	360
Db	301	GCTTCGTGGCGCTCTTCGCGGTGCGCTCTCTCTTCCTCTTCCTCTTCCTCTTCCTCTTC	360
Qy	361	CTTCCGCCATGAACCTGGCTTCTTGGCGCGCGACAGCTTCAGAAAGTGGCGCTCTCG	420
Db	361	CTTTCGCCATGAACCTGGCTTCTTGGCGCGCGACAGCTTCAGAAAGTGGCGCTCTCG	420
Qy	421	CGCTGTGGCGCTGGCGCTCCGCGGCTCTCTCTCCGCGCGCGCTCGGCTCGACTGGA	480
Db	421	CGCTGTGGCGCTGGCGCTCCGCGGCTCTCTCTCCGCGCGCGCTCGGCTCGACTGGA	480
Qy	481	GCATCAGCTCTTCTCCCTCTCCAGCTCCCAACACGCTCGTCACTGGGCATCCCGCTCG	540
Db	481	GCATCAGCTCTTCTCCCTCTCCAGCTCCCAACACGCTCGTCACTGGGCATCCCGCTCG	540
Qy	541	TGGAGGCACTGACGGCGGTCGTTCGGCGGCAAGCTCATGGTCCAGTTCGTCTCTCTCC	600
Db	541	TGGAGGCACTGACGGCGGTCGTTCGGCGGCAAGCTCATGGTCCAGTTCGTCTCTCTCC	600
Qy	601	AGTGATCATCTGGTACAGCTCATGCTCTTCTCTTCGAGTACCGCGCGCGCGCGCG	660
Db	601	AGTGATCATCTGGTACAGCTCATGCTCTTCTCTTCGAGTACCGCGCGCGCGCGCGCG	660
Qy	661	TCGTCTCGACAGTTCCCGACGGCGCGCGGTCCATGCTCTCTTCGGGTGACTCT	720
Db	661	TCGTCTCGACAGTTCCCGACGGCGCGCGGTCCATGCTCTCTTCGGGTGACTCT	720
Qy	721	CCGAGCTCGTCTCGCTCGCCAGGGGAGCGTCTGAGCTCGAGCGCGAGCCCGAGCGGTG	780
Db	721	CCGAGCTCGTCTCGCTCGCCAGGGGAGCGTCTGAGCTCGAGCGCGAGCCCGAGCGGTG	780
Qy	781	CCGGCGCGCGCGCTCTCTCTCCGCGCGGGAGCGCGCGCGGTCGCGTCAACGTCG	840
Db	781	CCGGCGCGCGCGCTCTCTCTCCGCGCGGGAGCGCGCGCGGTCGCGTCAACGTCG	840
Qy	841	GCAAGTCCACCACTCGCGCTCCGAGGCGCGTGTCTGCACTCCGACACCATGC	900
Db	841	GCAAGTCCACCACTCGCGCTCCGAGGCGCGTGTCTGCACTCCGACACCATGC	900
Qy	901	AGCCCCGTGTCCAACTCTTCGGCGTGGAGATCTACTCGCTGAGTGTGTCGGGCAAC	960
Db	901	AGCCCCGTGTCCAACTCTTCGGCGTGGAGATCTACTCGCTGAGTGTGTCGGGCAAC	960

Qy	961	CAACCCCGCGCGGGTCCAGTTCACACACGCCGACTTCTTTCAACATCGTTCGGCGCGCGCG	1020
Db	961	CCACCCCGCGCGGGTCCAGTTCACACGCCGACTTCTTTCAACATCGTTCGGCGCGCGCGCG	1020
Qy	1021	CCAAGGAGCGCGAGGAGCGCGCGGGGACAGAGAGAAAGGCGCATGCGCGCGCGCGCGCG	1080
Db	1021	CCAAGGAGCGCGAGGAGCGCGCGGGGACAGAGAGAAAGGCGCATGCGCGCGCGCGCGCGCG	1080
Qy	1081	GAGGACACTCGCCGACGCGCAGCGCGTTCGCGCTCGCGCCGCGGCAAGAGAAAGACCTGCACA	1140
Db	1081	GAGGACACTCGCCGCGCAGCGCGCGTTCGCGCTCGCGCCGCGGCAAGAGAAAGACCTGCACA	1140
Qy	1141	TGTCGCTGTGAGCTCCAGCGCTTCGCGCGTGTTCGAGCGCGCGCGTTCGAGCTCTTTCG	1200
Db	1141	TGTCGCTGTGAGCTCCAGCGCTTCGCGCGTGTTCGAGCGCGCGCGTTCGAGCTCTTTCG	1200
Qy	1201	GCSCCGCGCGCTGACCACTACCGCAGCTCTCCCAAAGGAGCCAGGCTCTACGACGAGT	1260
Db	1201	GCSCCGCGCGCTGACCACTACCGCAGCTCTCCCAAAGGAGCCAGGCTCTACGACGAGT	1260
Qy	1261	ACGGCGCGACGACTACAGCAGCAGACGAAAGACGGGAGCGCGCGCGGACGAGGGCG	1320
Db	1261	ACGGCGCGACGACTACAGCAGCAGACGAAAGACGGGAGCGCGCGCGGACGAGGGCG	1320
Qy	1321	GGCGGAGCTGTGAGAGCTGAGCTCGACGCGGCGAGCTGTATCCCAAAGGACGACG	1380
Db	1321	GGCGGAGCTGTGAGAGCTGAGCTCGACGCGGCGAGCTGTATCCCAAAGGACGACG	1380
Qy	1381	GCAGGGAGGGCGCGCGGTGGCGATCCGCGCGGAGCGGTGATGACCGCGCTCATCC	1440
Db	1381	GCAGGGAGGGCGCGCGGTGGCGATCCGCGCGGAGCGGTGATGACCGCGCTCATCC	1440
Qy	1441	TCATCATGTTGTGAGGAGCTGATCCGGAACCCCAACACTACTCCAGAGCTCATCGCGG	1500
Db	1441	TCATCATGTTGTGAGGAGCTGATCCGGAACCCCAACACTACTCCAGAGCTCATCGCGG	1500
Qy	1501	TCGTCGTGTCCTGGTTCCTACAGTGGGGCATTCGAGATTCGACGCAATATCGCCCGGT	1560
Db	1501	TCGTCGTGTCCTGGTTCCTACAGTGGGGCATTCGAGATTCGACGCAATATCGCCCGGT	1560
Qy	1561	CGATTTCCATCTGTGCGAGCGCGGTCTCGGGATGGCCATGTTTCAGGCTTAGGCCCTGTTC	1620
Db	1561	CGATTTCCATCTGTGCGAGCGCGGTCTCGGGATGGCCATGTTTCAGGCTTAGGCCCTGTTC	1620
Qy	1621	TGGCGCTGCAGCGAGGATCATCGCTGCGGGAACAAGCTTGGCGGCCATCGCGATGGCG	1680
Db	1621	TGGCGCTGCAGCGAGGATCATCGCTGCGGGAACAAGCTTGGCGGCCATCGCGATGGCG	1680
Qy	1681	TCGCGTTCGTCGAGGCGCGCGGTATGCGCGCGCGCTCCATCGCGCGTTCGCTTCGCGG	1740
Db	1681	TCGCGTTCGTCGAGGCGCGCGGTATGCGCGCGCGCTCCATCGCGCGTTCGCTTCGCGG	1740
Qy	1741	GGCTTCCTCTCCATCGCCATCGTCAGAGCTGTCTGCCTCAGGCGATCGTGCCTTCG	1800
Db	1741	GGCTTCCTCTCCATCGCCATCGTCAGAGCTGTCTGCCTCAGGCGATCGTGCCTTCG	1800
Qy	1801	TGTTCCGCAAGAGGTACGGCGTTTCATCCCGACATCTCTGAGCACAGCGTATGGTCCAATAA	1860
Db	1801	TGTTCCGCAAGAGGTACGGCGTTTCATCCCGACATCTCTGAGCACAGCGTATGGTCCAATAA	1860
Qy	1861	CATCGCATGGTTTCATCACTTGCATAGTTAAACGGGAAACAAAGACAGACCAATCCGATG	1920
Db	1861	CATCGCATGGTTTCATCACTTGCATAGTTAAACGGGAAACAAAGACAGACCAATCCGATG	1920
Qy	1921	ACGACGCACTGAATTCATTAATGATTTACTAAATGATGGTGTGTTTCATGCAGTGCAGTC	1980
Db	1921	ACGACGCACTGAATTCATTAATGATTTACTAAATGATGGTGTGTTTCATGCAGTGCAGTC	1980
Qy	1981	AAAGAACACTAATAAGCATGTACTAGGACGACATCAGGATGATGATTCCTTGTGTTTC	2040
Db	1981	AAAGAACACTAATAAGCATGTACTAGGACGACATCAGGATGATGATTCCTTGTGTTTC	2040
Qy	2041	TCCTGCAAACTTCGATTTCTTACTACAGCTGCTCCTCACTCAATCCATCAGATGATCA	2100

2041	Db	TCCTGCAATCTCGATTTCTTACTACACAGTGTGCTTCTACTCATCCATCCAGATGATCA	2100
2101	QY	TACAAACACTACTGATGATCTTTTTTTGTGATCTGCTGCAGCGTGATCTTCGGGATGCT	2160
2101	Db	TACAAACACTACTGATGATCTTTTTTTGTGATCTGCTGCAGCGTGATCTTCGGGATGCT	2160
2161	QY	CATCGCTCTGCCCATCACCCCTGGTCTACTACATCTTCTCGGGGCTGTGAGCCCTCTCTCGC	2220
2161	Db	CATCGCTCTGCCCATCACCCCTGGTCTACTACATCTTCTCGGGGCTGTGAGCCCTCTCTCGC	2220
2221	QY	TCGCTTCTTCAGCGTGCAGAGGCGCCATGCTGTGTGTATCGGCCCCACATGAAATTTTC	2280
2221	Db	TCGCTTCTTCAGCGTGCAGAGGCGCCATGCTGTGTGTATCGGCCCCACATGAAATTTTC	2280
2281	QY	TGATGACATTAGCGATTACTATTAGCTTAGCGAAGAATGATGATGGTGTGGCCTGT	2340
2281	Db	TGATGACATTAGCGATTACTATTAGCTTAGCGAAGAATGATGATGGTGTGGCCTGT	2340
2341	QY	CGGACTGGGGGAGTCAGACAGACCCCTCGAACAAAGTTTCITTTGGCTTCCTGCTCC	2400
2341	Db	CGGACTGGGGGAGTCAGACAGACCCCTCGAACAAAGTTTCITTTGGCTTCCTGCTCC	2400
2401	QY	GTCAAGAACAAAGTTTGGCTTTTGGCATCGGCACCTCGAAGCACAGCAGCAGCAGCAGC	2460
2401	Db	GTCAAGAACAAAGTTTGGCTTTTGGCATCGGCACCTCGAAGCACAGCAGCAGCAGCAGC	2460
2461	QY	ATCATTCATGAGATGATCTCTCGAATCTTAGAGCTAGCGAAGGCAATATTAAGATA	2520
2461	Db	ATCATTCATGAGATGATCTCTCGAATCTTAGAGCTAGCGAAGGCAATATTAAGATA	2520
2521	QY	CCACAGGCAATGGAATCAACAAAGCTTCATGCGAGCGGCTATCATATCAAGGAACACA	2580
2521	Db	CCACAGGCAATGGAATCAACAAAGCTTCATGCGAGCGGCTATCATATCAAGGAACACA	2580
2581	QY	TGCAGAAATCAACCGAGTCTAGTCGCAATGGCTTCCTCTTTTCTTTCTTCGCGAAA	2640
2581	Db	TGCAGAAATCAACCGAGTCTAGTCGCAATGGCTTCCTCTTTTCTTTCTTCGCGAAA	2640
2641	QY	GGGTTCCTAGACTGATTAAGGATTCCAATAGCATCTCTGGATTCGATTCCTTCCGAG	2700
2641	Db	GGGPTTCCTAGACTGATTAAGGATTCCAATAGCATCTCTGGATTCGATTCCTTCCGAG	2700
2701	QY	ACAAATTTCTGGCTTTTGTAGAAAAATCCTCTCGTTGAAAAAATAAAAAAAAAAAAAA	2760
2701	Db	ACAAATTTCTGGCTTTTGTAGAAAAATCCTCTCGTTGAAAAAATAAAAAAAAAAAAAA	2760
2761	QY	AAAAAAAAAG 2769	
2761	Db	AAAAAAAAAG 2769	

RESULT 2  
ADA70391  
ID ADA70391 standard; DNA: 1926 BP.

AC	ADA70391;	
NC		
XX		
XX		
DT	20-NOV-2003	(first entry)
XX		
XX		
XX		
XX		
XX		
XX		
XX		
XX		
KW		
KW		

XX  
OS

Oryza sativa.

XX PN WO2003000898-A1.

XX  
PD  
03-JAN-2003.XX  
DE 22-JTN-2001;

XX

PR	22-JUN-2001; 2001WO-IB001105.
XX	
PA	(SYGN ) SYNGENTA PARTICIPATIONS AG.
PI	
XX	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;
PI	Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;
XX	
DR	WPI; 2003-175290/17.
XX	
PT	Identifying at least one gene involved in plant resistance or response to
PT	pathogenic infection for conferring resistance or tolerance to a plant
PT	bacterial, fungal or viral infection by determining or detecting plant
PT	gene expression.
XX	
XX	
PS	Claim 6; SEQ ID NO 3714; 899pp; English.
XX	
CC	The present invention relates to a method (M1) for identifying genes
CC	involved in plant resistance or response to pathogenic infection. M1
CC	comprises identifying a gene whose expression is significantly altered in
CC	the incompatible interaction of plant gene expression relative to
CC	expression of the gene in an uninfected plant, in a mutant plant that
CC	does not express a gene associated with response to pathogenic infection,
CC	or in a corresponding incompatible or compatible interaction. (M1) is
CC	useful for conferring resistance to resistance or tolerance to a plant to
CC	bacterial, fungal or viral infection. The present sequence was used to
CC	illustrate the invention.
XX	
XX	Sequence 1926 BP; 328 A; 612 C; 556 G; 427 T; 0 U; 3 Other;

	Query Match.	30.3%;	Score 838.6;	DB 7;	Length 1926;
	Best Local Similarity	76.5%;	Pred. No. 5.3e-125;		
	Matches 1244;	Conservative	3;	Mismatches 257;	Indels 123; Gaps 13;
Qy	171	ATGATCACCGCGGTGACCTTACCACAGTGCTCAGCGCGGTGTGTCGTGTACGTGGCC	230		
Dd	1	ATGATCACGGTGGTGACCTGTACCACGCTCTCAGCGCGGTGTGTCGTGTACGTGGCG	60		
Qy	231	ATGACGTTGCCTACGGCTCCGCTCCGCTGGGTGGGCACTTTCACGCCGACCACTGCTGCC	290		
Dd	61	ATGACGTTGCCTACGGCTCCGTCGGGTGGGTGGGCACTTTCCTCCGACCACTGCTGCC	120		
Qy	291	GGGATCAACCGCTTCGTGGCGCTTTCGCGTGCGGCTCTCTCTCTCTTCCACTTCATCTCC	350		
Dd	121	GGCATCAACCGCTTCGTGGCGCTTTCGCGTGCGGCTCTCTCTCTCTTCCACTTCATCTCC	180		
Qy	351	ACCAAAGCACCCCTTCGGCATGAACCTTGCCTTCTGGCCGCGACACGCTGCAGAAGGTG	410		
Dd	181	ACCAACAACCCCTTCGGCATGAACCTTCGGCTTCTCGCCGCGACACGCTCCAGAAGCTC	240		
Qy	411	GCGCTCTTCGCGTGTGGCGCTTTCGCGGCGCTCTCTCTCCGCGCGCGCTCGGG	470		
Dd	241	ATCGTCTCTCCCTCTCTCGGSCF-----CTGGTGGCGCTCTCTCCGCGCGGCTCM---	291		
Qy	471	CTCGACTGGAGCATCAGCTCTTCTCCCTCTCCACGCTCCGCCAACACGCTCGTCATGGCC	530		
Dd	292	CTCGACTGGCTCATCACCTCTTCTCCCTCTCCACCCCTCCCAACACCTCGTCATGGCC	351		
Qy	531	ATCCGCGTGTGGAGGATGTACGGCGCTCGTGG-----CGGCG	572		
Dd	352	ATCCGCGTGTCAAGGGGATGTACGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	411		
Qy	573	ACGCTCATGTTCCAGTTCGTGCTCTCCAGTGCATCATCTGGTACACGCTCATGCTCTTC	632		
Dd	412	AGCCTCATGTTGCAGATCGTCTGCTCCAGTGCATCATCTGGTACACGCTCATGCTGTTCT	471		
Qy	633	CTCTTCAGTACCGCGCGCGCGCGCGCTCGTCTCTCGACCAAGTTCGCCGACGGCGCGGCC	692		
Dd	472	CTCTTCAGTACCGTGGCGCGCGCTCTTGTCTATGAGCAGTTCCCCGGA---CACGCC	528		
Qy	693	GCGTCCAATCGTCTCTTCCTTCGCGTGCATCTCCGACGTCTCTCGCTCGCAGGGGGAGTGC	752		
Dd	529	GCTTCCAATCGTCTCTTCCTTCGCGGTGGATTCCGACGCTCTCTCGCTCGCGCGCGGGTGGT	588		







PA (WHED) WHITEHEAD INST BIOMEDICAL RES.  
XX Luschnig C, Gaxiola RA, Grisafi P, Fink GR;  
XX WPI; 2000-086979/07.  
DR P-PSDB; AAY44265.  
XX  
XX DNA encoding a root-specific auxin transport protein, used to develop  
XX transgenic plants with increased resistance to herbicides.  
XX  
XX Claim 1; Fig 8; 55pp; English.  
XX  
XX The present sequence is a cDNA encoding REH1, a rice homologue of EIR1  
XX which functions as a root-specific auxin transport (efflux) carrier  
XX protein involved in gravitropism. The sequence is obtained from a rice  
XX EST derived from root-specific cDNA. The sequence is used for producing  
XX genetically engineered plants with greater resistance to auxin-based  
XX herbicides and auxin transport inhibitors in combination with a second  
XX herbicide. It can also be used to enhance transport of auxin in plant  
XX roots, produce transgenic plants which exhibit altered auxin homeostasis  
XX and mutant plants in which the roots are agravitropic and have reduced  
XX sensitivity to ethylene  
XX  
XX Sequence 2374 BP; 471 A; 736 C; 733 G; 434 T; 0 U; 0 Other;  
XX  
XX Query Match 24.2%; Score 671.2; DB 3; Length 2374;  
XX Best Local Similarity 66.8%; Pred. NO. 3e-98;  
XX Matches 1196; Conservative 0; Mismatches 423; Indels 171; Gaps 10;  
XX  
XX 165 GCGGCCATGATCACCGCGTGTGACCTTACACGCTGTGACGGGTGTGTCCTGTAC 224  
XX Db 152 GCGAAGATGATTACCGCGCGGACCTTACACGCTGTGACGGGTGTGTCCTGTAC 211  
XX  
XX 225 GTGGCCATGACGCTGGGTACGGTTCCTGCGTGGTGGGCGATCTTACGCGGACGAG 284  
XX Db 212 GTGGCGATGATCTGGCGTACGGGTGCGTGAAGTGGTGGCGATCTTACGCGGACGAG 271  
XX  
XX 285 TGCTCCGGGATCAACCGCTCTGTGGCGCTCTTCCGCGTGGCGCTCTCTCTTCCATTC 344  
XX Db 272 TGCTCCGGGATCAACCGCTTGTGGCGCTCTTGGCGCTGTGGCGCTGTGGTTCATCTTC 331  
XX  
XX 345 ATCTCCACCAAGACCCCTTCCGCAATGAACCTTGGCTTCTTGGCGCGGACACGCTCGAG 404  
XX Db 332 ATCTCCACCAACACCCGTACACGATGAACCTTCCGGTTTCATCGCGCGGACACGCTCGAG 391  
XX  
XX 405 AAGTGGCGCTCTCGCGCTGTGCGCTGTGCGCTTCCGCGGCTCTCTCTCCCGCGCGG 464  
XX Db 392 AAGTGTATGGTGTGGCGATGCTACGCGGTGAGACCACTCAGC-----CGCGG 442  
XX  
XX 465 CTCGGGCTCGAGTGGAGATCACGCTCTTCTCCCTCTCCAGCTTCCCAACACGCTCGTC 524  
XX Db 443 GGGAGCCTCGAGTGGACCATCACGCTCTTCTCCCTCTCCAGCTTCCCAACACGCTCGTC 502  
XX  
XX 525 ATGGGATCCCGCTCTCGGAGGATGTACGGCGGTGTGCGCGGCGGACGCTCATGCTC 584  
XX Db 503 ATGGGATCCCGCTTGTCTCAAGGGCATGTACGGGAGTTCT---CGGGAGCCTCATGCTG 559  
XX  
XX 585 CAGGTCTGCTCTCCAGTGCATCATCTGTACAGCTCATGCTTCTTCTCTCGAGTAC 644  
XX Db 560 CAGATCTGCTCTCGAGTGCATCATCTGTACAGCTCATGCTTCTTCTTCTTCTGAGTAC 619  
XX  
XX 645 CGCGCGCGCGCGCTGCTCTCGACCAAGTTTCCCGACGCGCGCGCGCTTCCATCGTC 704  
XX Db 620 CGCGCGCGCGCGATGCTCATCACGAGCAGTTCCTCGGA---CACCGCGCGCAACATCGCC 676  
XX  
XX 705 TCTTTCGGGTGATCTCGAGTGTCTGTCTGCTCGCAGGGGGAGCTGAGCTCGAGGCC 764  
XX Db 677 TCCATGTGTGACCCCGACGCTGTGTGTGTGAGCGGAGGAGGCCATCGAGACG 736  
XX  
XX 765 GAGCCCGACGCGCTCGCGCGCGCGCGCTCTCTCTCCCGCGGGGAGCGCGCGCGG 824  
XX Db 737 GAGACGGA-----GCTAAGAGGAGACGGCAGG 763

QY 825 GTCCGGTCAACGTCGCGAAGTCCACAGCTCCGAGCGCGCTGCTCGCACTC- 883  
Db 764 ATACAGTCAACGTCGCGCTCCACAGCGTCTCGTTCGGAATCTACTCCCGCGCTCC 823  
QY 884 --GCATCCAGACCATGAGCCCGTGTGTCAACCTCTCCGCGCTGAGATCTACTCG 941  
Db 824 ATGGGCTTCTCCAGCACACGCGCGCGGAGCACTTCCACCAACCGGAGATCTACTCG 883  
QY 942 CTGCACTGTGTCGCGCAACCCACCCCGCGCGGTCCAGTTCAGTTCACACCGCACTTCTTC 1001  
Db 884 CTGCACTGTGTCGCGCAACCCAGCGCGGAGGTTCAGGTTCAACACACCGCACTTCTAC 943  
QY 1002 AACATCTGTCGCGCCCGCCCAAGGAGCGGAGGAGCGCGGGGAGCAGAGAGAGGGC 1061  
Db 944 TCCATGTCGCGCGCAGCTTCCAACTTCGCGCGCGCGAGCGCTTCGCGCTCCGCAACGCG 1003  
QY 1062 GC----- 1063  
Db 1004 GCACGCGCGCGCCCTTCACTAGAGGAGCAGCGGTCCAGGCCCAAGTACCGGTCCCG 1063  
QY 1064 -----ATGCGCGCGCGCGCGGAGGAC-ATCGCGCGACCGCGCGCTTCG- 1111  
Db 1064 GCGTCGAATGCGCGCGCCATGCGCGGCCACTACCGCGCGCGGAAACCGCGCGCTGCTCG 1123  
QY 1112 -----CGTCCGCGCAAGAGGAGCACTGCA 1139  
Db 1124 GCGCCCAAGGCGCCCAAGAGGCGGCCACGAAACGGGCAAGGCGGAGGACCTCCAC 1183  
QY 1140 ATCTCTGTCGAGCTTCCAGCGCTCCGCGCTGTCGAGCGCGCGCGCTGACGCTTTC 1199  
Db 1184 ATGTTGTTGAGCTTCCAGCGCTTCCGCGCTGCGCGTTCG-----ACGCTTTC 1228  
QY 1200 GCGCGCGCGCGCTGACCATGCGGAGCTCTCGCCAAAGGAGCCAGGCTTACGAGAG 1259  
Db 1229 GCGCGCGCGCGCGACACTACAAACGCGCGCGGAGTCAAGTCCCGCGCAAAATGGAT 1288  
QY 1260 TAGCGCGCGACACTACAGCAGCAGAGAGAGAGAGCGCGCGCGCGCGCAAGAGGCG 1319  
Db 1289 GAGCGAAGAGCAGGAGGACTACCTGAGAGCGGAGCTTTCAGCTTCGGAACAGGCGC 1348  
QY 1320 G---GGCCGACGCTGTGAAAGTGGGTCCAACTCGACGCGCGAGCTGTACCCCAAGGAC 1376  
Db 1349 GTCATGACAGGAGACCGGAGGACGAGAGGCGCGCGCGCGCGCGCGCGCGCGCG 1408  
QY 1377 GAGCGGAGGAGGCGCGCGGTGGCGATGCGCGCGCGCGAGCGTGTATGACGCGGCTC 1436  
Db 1409 CCAGCAAGCCATGCGCGCGCGCGAGCGGATGCGCGCGAGCGGTGATGACCGGCTTC 1468  
QY 1437 ATCTCATCATGTCGTGAGGAGTGTATCCGGAACCCCAACACCTTATCCAGCTCATC 1496  
Db 1469 ATCTCATCATGTCGTGAGGAGTGTATCCGCAACCCCAACACCTTATCCAGCTCATC 1528  
QY 1497 GCGCTGTCGTGTCCTGCTTCTTACAGTGGGCGATCGAGATGCGAGCATCATCGCC 1556  
Db 1529 GCGCTCATGTCGTGTCCTGCTTCTGCTTTCAGTGGAACTTCAGATGCGCGCATCTCTG 1588  
QY 1557 CGGTGATTTGATCTCTGAGCGCGGTCTCGGATGCGCATGTCAGCTTACGCTTACGCTG 1616  
Db 1589 AATATCATCTGATCTGTGAGCGCGGCTCGGATGCGCATGCGCTTTCAGTCTCGGCTG 1648  
QY 1617 TCCATGCGCTGAGCGGAGGATCATCGCTGCGGGAACAGCTGCGGCGCATCGGATG 1676  
Db 1649 TCCATGCGCTGAGCGGCGCATCATCTGCGTGGGAAACAGGTGCGGAGCTACGCGCATG 1708  
QY 1677 GCGCTGCGGTTCGTTCGAGCGCCCGCGGTCTGCGCGCGCTTCCATCGCGCTCGGTCTG 1736  
Db 1709 GCGGTGCGGTTCGTTCGCGCGCGCGCGGTGATGCGCGCGCTTCTTTCGCGCTCGGACTC 1768  
QY 1737 GCGCGGCTCTCTCTCATATCGCATCTGTCAGGCTGCTCTGCTCAGGAGTCTGCGCG 1796  
Db 1769 CGTGGCAGCTCTCTGACGCTGCGCATTTGTCAGGAGCTCTGCGCGCGGAGCTTTCGCC 1828  
QY 1797 TTGCTGTTCGCCAAGGAGTACGCGGTTCATCCGACATCTCTGAGCACAGC 1846

Db 1829 TTCGTCTTCCCAAGGAGTACACGCTGCACCTTAGCATTTCTCAGCACAGC 1878  
RESULT 6  
ID AAA94733 standard; DNA; 2293 BP.  
XX AC AAA94733;  
XX DT 02-FEB-2001 (first entry)  
XX DE Wheat auxin transport protein clone wdk1c.pk008.g1 DNA sequence.  
XX KW Auxin transport protein; wheat; root development; gene mapping;  
XX KW plant breeding; herbicide; ss.  
XX OS Triticum aestivum.  
XX PN WC200068389-A2.  
XX XX  
XX PD 16-NOV-2000.  
XX PF 03-MAY-2000; 2000MO-US012061.  
XX PR 07-MAY-1999; 99US-0133040P.  
XX PA (DUPO) DU PONT DE NEMOURS & CO E I.  
XX PA (PION-) PIONEER HI-BRED INT INC.  
XX XX  
XX PI Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;  
XX XX  
XX DR WPI; 2000-687647/67.  
XX DR P-PSDB; AAB26946.  
XX PT New nucleic acid sequences encoding new auxin transport proteins, useful  
XX PT for modulating root growth of plants and to screen for herbicides.  
XX XX  
XX PS Claim 2; Page 84; 94pp; English.  
XX CC Auxins are plant hormones that influence plant behaviour and development  
XX CC e.g. vascular tissue differentiation, apical development, tropic  
XX CC responses and organ (e.g. flower, leaf) development. The present  
XX CC invention relates to corn auxin transport protein coding sequences and  
XX CC proteins. The present sequence is one such coding sequence. This sequence  
XX CC may be used to modulate root development, e.g. to produce a more robust  
XX CC root system, alter root angle or redirect root growth. Also, the present  
XX CC sequence may be useful for gene mapping (e.g. for plant breeding) and to  
XX CC identify loss of function mutants. The protein encoded by the present  
XX CC sequence may be useful for raising specific antibodies, for the detection  
XX CC of auxin transport proteins and to design and/or identify specific  
XX CC inhibitors of auxin transport proteins, potentially useful as herbicides  
XX SQ Sequence 2293 BP; 462 A; 713 C; 683 G; 435 T; 0 U; 0 Other;  
Query Match 22.6%; Score 626.8; DB 3; Length 2293;  
Best Local Similarity 64.5%; Pred. No. 3.8e-91;  
Matches 1199; Conservative 0; Mismatches 467; Indels 192; Gaps 10;  
QY 164 CGCGGCCATGATCACCAGCGTGCACCTCTACACGCTGCTGACGGCGGTGGTCCCGTGA 223  
Db 67 CGTCCGATGATCACCAGGGAAGACATCTACGACGTGCTGGCGGGGTGGTCCCGTGA 126  
QY 224 CGTGGCCATGACCTGGCGTACGGCTCCGCTCGCTGGTGGCGCATCTTCAACCGCGACCA 283  
Db 127 CGTGGCCATGTTTCATGGCGTACGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 186  
QY 284 GTGCTCCGGGATCAACCGCTTCTGGCGGTCTTTCGGCGGTGGCGGTCTCTCTCTCCACTT 343  
Db 187 GTGCTCCGGGATCAACCGCTTCTGGCGGTCTTTCGGCGGTGGCGGTCTCTCTCTCCACTT 246  
QY 344 CATCTCCACCAACGACCCCTTCCCATGAACCTGCGCTTCTGGCGGTGGCGGTGGCGGTGGCG 403

Db 247 CATCTCCACCAACGACCCCTTAGCCATGAGTACCGCTTCTTGGCCGCCGACCTGCTGCA 306  
QY 404 GAAGGTGGCCGTCTCTGCGCTGCTGGCGCTTCCCGCGGCTTCTCTTCCCGCGCGC 463  
Db 307 GAAGCTCGTCACTCTGCGCGCTTCCCGCTTGGCAACAAGTGTCTCTCCCGCTACCGGTG 366  
QY 464 GCTCGG-----GCTCGACTGGAGCATACGCTCTTCTC 496  
Db 367 CCGCGCGGCACGAGGCGCGGAGGCTCGTCTGAGACTGGACCATACGCTCTTCTC 426  
QY 497 CCTCTCACGCTCCCAACACGCTCGTCAATGGGATCCCGCTGCTGCGAGGCAATGACGG 556  
Db 427 CTTGGCAGCTGCGCAACACGCTGATGGGATCCCGCTGCTGCGCGCCATGACGG 486  
QY 557 CGGTCGTGGCGGCACGCTCATGTCAGGTCGCTCTCTCCAGTGCATCATCTCTGTA 616  
Db 487 CGACTTCTCG---GGTCTCGTCTGTCAGATGCTGTCGAGAGCTCATCTCTGTA 543  
QY 617 CACGCTCATGCTCTTCTCTTTCGAGTACCGCGCGCGCGCGCTCTCTCTCGACCAAGTT 676  
Db 544 CACGCTCATGCTCTTCTCTTTCGAGTACCGCGCGCGCGCGCTCATCTCTCGAGCAGTT 603  
QY 677 CCGCAGCGCGCGCGCGCGCTCATGTCAGGCGCGGAGCGCGCTGCGCGCGCGCGCGCT 736  
Db 604 CCGCGCGGAGCTCGGCGCGGAGCTCGCTCTCTTCCGCTCGACTCGCGCTCTCTGCT 663  
QY 737 CGCAGGCGGAGCGTCTGAGCTCGAGGCGGAGCGCGCTGCGCGCGCGCGCGCGCGCT 796  
Db 564 CAACGGGCGGAGCGCTGCAACCGGAGCTGGCGCGGAGCGCGCGCGCTCGACGT 723  
QY 797 CTCCTCCCGCGCGGAGCGCGCGGCTGCGCTCACACGCGCGCGCGCGCGCGCGCTCG 856  
Db 724 CGTCACTCCCGCGGCTCGCGCTCGGCGCTCAACGCGCGCGCGCGCGCGCGCTCGCG 783  
QY 857 GCGTCTCGAGCGCGGCTGCTCGCACTCGCACTCCAGACCATGCAAGCGCGCGCTGTCCAA 916  
Db 784 GATCTACGCTGGCGG-----TCAACGCCATGACGCGCGCGCGCGCGCGCT 828  
QY 917 CCTCTCCGGCTGGAGATCTACTCGTGTGAGTCTGTCGCAACCCACCGCGCGCGGCT 976  
Db 829 CCTCACGGGCTGGAGATCTACTCGCTGCAAGCTGCGCGGAGCGCGCGCGCGCGCT 888  
QY 977 CAGCTTCAACCGCGCGCTTCTTCAACATCTGCGG----- 1014  
Db 889 CAGCTTCAACCGCTCCGACTTCTACTTCCATGTTCAACGGAGCAAGTGGCTAGTCCAA 948  
QY 1015 -----CGCGCGCCAAAGGAGCGGAGCGCGCGG-----GACGAGAGAA 1057  
Db 949 GGGCCAGCGCGCGCTCGCGGAGTGTGTGTGCGCGCGCGCGCGCGCGCGCTCGACGAGGT 1008  
QY 1058 GGGCGCATCGCGCGCGCGCGGAGACACTCGCGCAGCGCGCGCGCTGCGCGT-- 1115  
Db 1009 GGGCAACAAGTTCAAGGGCGGAGGCGGCTGCGCTTACCGCGCGCGCGCGCGCGAT 1068  
QY 1116 -----CGGCGCAAGAGGAGGA 1132  
Db 1069 GATGATCGCGCGCGCACGGAAGAGCTTGGGGTTCCAACTCAAACTCGAACAGGA 1128  
QY 1133 CTGTCATGCTGTGTGAGCTCCAGCGCTCGCGGTGTCGAGCGCGCGCGCGCTGCA 1192  
Db 1129 GCTGCAATGTTGTGTGAGCTCCAGCGCTCGCGGTGTCGAGGCGCGCGCGCGCGCG 1188  
QY 1193 ---GCTCTTGGCGCGCGCGCTGACCATGTCGAGTCTCTCGCCAAAGAGCGCGCG 1249  
Db 1189 CGCGCTCAACACCGCGCTCCACCGACTTTCGCGCGCGCGCGCGCGCGCGCGCGCG 1248  
QY 1250 CTACGAGAGTACGGGCGCGAGCACTACAGCAGCAG-----GACGAGAA 1294  
Db 1249 ACAGAGCG 1308  
QY 1295 CGGAGCG 1333  
Db 1309 CGCAGGAGCGCGCGCTGGAGTGGAGTGGAGCGGCGCATGATGAAGAGCGCGCGCG 1368



Db 776 TGGCGGGCTCGAGCTTCGGCGCTCCGAGCTCTACTCGTCAATCGTCGGGGCCCAA 835  
Qy 1042 CGGGGACAGGAGAGGCGCATCGCGGGCGGGGAGGACACTCGCGGACGCGC 1101  
Db 836 CCCCAGGAGCTCCACTTCGACGAGCACTCGGACGCGCGGAAACACCGGCAACA 895  
Qy 1102 AGGCGGTGCCGTCGCGGCAAGAGAGGAGCACTGCAATGCTGCTGTGAGCTCCAGC 1161  
Db 896 CCACGGGGCACTCAACACGATGCGCAAGAGCTCCACATGTTCTGTGTGAGCTCGAGC 955  
Qy 1162 CTTCCGCGCTGTCGAGCGCGCGCGCTGCACTCTTC---GGCGCGCGCGCTGACC 1218  
Db 956 CGTCTCCCGTCTCAGAAGTCAGCGGCTGCTGTGTTGAGTGGCGGCGCGCGCGG 1015  
Qy 1219 ATGCCGACGCTCTCGCAAGAGAGGCCA----- 1246  
Db 1016 CTCTGACGTGCGCGCAAGGAATCCACATGTCATCCCGCGACCTGCGCGAGACA 1075  
Qy 1247 -----GGCTACGAGC 1257  
Db 1076 AGGCTCAGGCAAGAGACGAGGAGTACGGCGAGTGGGTGGCGCGCGCGCG 1135  
Qy 1258 AGTACGGGGCGACGACTACAGCAGCAGGACGAAGAACGGAGCGCGCGCGCAAG 1317  
Db 1136 GAGAGAACTTCACTTCGGAGGCGGCAAGCGTGGACGCGCGGAGGAGTAGACGAG 1195  
Qy 1318 GCGGGC-----GACGCTGTGAAGCTGGGTCCAACTCGAGCGCGAGCTGTACC 1368  
Db 1196 AGCGGCTTTCCTGTCAGCGGCTGACGAAGATGGGTGCGAGCTCGACGCGGAGCTGCACC 1255  
Qy 1369 CCAAG-----GACGCGCGGAGGAGGCGCGCGCGG 1401  
Db 1256 CGAAGTCTGTCAGCTCGACGACGAAACGCGCGCGCGCGCGCGCGCGAGT 1315  
Qy 1402 TGGCGATGCGCGCGGAGGCTGATGACGCGGCTCATCTCATGATGTTGAGGAGAC 1461  
Db 1316 ACCAAATGCGCGGCGGAGGCTGATGACACGCGCTCATCTCATGTTGTTGGCGCAAGC 1375  
Qy 1462 TGATCGGAGACCCCAACACTACTTCCAGCTCATCGGCTGCTGCTGCTGCTGCTCT 1521  
Db 1376 TCATCCGACCCCAACACTTCTCAGCTCTCGGCTCTCGGCTGCTGCTGCTGCT 1435  
Qy 1522 ACAGTGGGCTCGAGATGCCAGCATATGCCCGCTCGATTCGATTCCTGTGCGAGC 1581  
Db 1436 TCCGTTGGCAGCTTCCATGCCAGCAATCGTCGAGAAGTCCATCTCCATCTCTCGAGC 1495  
Qy 1582 CGGCTCTCGGATGGCATGTTCAAGCTAGGCTGTTTCATGGGCTGCGAGCGAGGATCA 1641  
Db 1496 CAGGCTTGGGATGGCCATGTTAGCTGGATGTTTCAAGGCTGCGAGCCAGCATCA 1555  
Qy 1642 TCGCTGCGGGAAACAAGCTGCGCGCCATCGCATGGGCGTTCGGTTTCGTCGAGGCGCG 1701  
Db 1556 TCGGTTGGCAATCAGCGCGCTGCTCTCCATGGCGCTTCGCTTCCTCGGCGCGCTG 1615  
Qy 1702 CGGTATGGCGCGGCTCCATCGCGTGGTCTGCGCGGCTGCTCCCTCCACATCGCCA 1761  
Db 1616 CCGTCATGGCGCGGCTCAATCGCATCGGATCCGCGGAGCGCTCTCTGACGTGCCCA 1675  
Qy 1762 TCGTCAAGGCTGCTCTGCTCAGGGATCGTGGCGTTTGGTTTCGCAAGGAGTACGGCG 1821  
Db 1676 TTGTTCAAGCGGCTTACACACAGGATGTTGCTTTGTTTGAAGAAATACAATG 1735  
Qy 1822 TTGATCCCGACATCTGAGCAGAGC 1847  
Db 1736 TCCACCGGCTATCTGAGCAGAGC 1761

RESULT 8  
ADA69396  
ID ADA69396 standard; DNA; 1845 BP.  
XX  
AC  
AC  
ADA69396;  
XX

DT 20-NOV-2003 (first entry)  
XX Rice gene, SEQ ID 2719.  
XX Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.  
XX Oryza sativa.  
XX WO2003000898-A1.  
XX 03-JAN-2003.  
XX 22-JUN-2001; 2001WO-IB001105.  
XX 22-JUN-2001; 2001WO-IB001105.  
XX (SYGN) SYNGENTA PARTICIPATIONS AG.  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.  
XX Identifying at least one gene involved in plant resistance or response to  
XX pathogenic infection for conferring resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection by determining or detecting plant  
XX gene expression.  
XX Claim 6; SEQ ID NO 2719; 899pp; English.  
XX The present invention relates to a method (M1) for identifying genes  
XX involved in plant resistance or response to pathogenic infection. M1  
XX comprises identifying a gene whose expression is significantly altered in  
XX the incompatible interaction of plant gene expression relative to  
XX expression of the gene in an uninfected plant, in a mutant plant that  
XX does not express a gene associated with response to pathogenic infection,  
XX or in a corresponding incompatible or compatible interaction. (M1) is  
XX useful for conferring resistance to resistance or tolerance to a plant to  
XX bacterial, fungal or viral infection. The present sequence was used to  
XX illustrate the invention.  
XX Sequence 1845 BP; 315 A; 628 C; 557 G; 345 T; 0 U; 0 Other;

Query Match 20.3%; Score 563; DB 7; Length 1845;  
Best Local Similarity 63.5%; Pred. No. 5.8e-81;  
Matches 1158; Conservative 0; Mismatches 430; Indels 237; Gaps 10;  
Qy 211 TGGTGGCGCTGTAGTGGCCATGACGCTGGCTACGCTCCGCTGGTGGCGCATCT 270  
Db 2 TGGTGGCGCTGTAGTGGCCATGATACTGGCTACGGTGGTGGCGCATCT 61  
Qy 271 TCAGCGCGGACCACTGTCTCGGGATCAACCGCTTCTGGGGCTTTCCGCGTGGCGTCC 330  
Db 62 TCAGCGCGGACCACTGTCTCGGGATCAACCGCTTCTGGGGCTTTCCGCGTGGCGTGC 121  
Qy 331 TCTCTTCCCTTCACTTCACTCCACCAACGACCCCTTCGCATGAACCTCGGCTTCTGGCGG 390  
Db 122 TGTGTTCCATTTCACTTCCACCAACCAACCGGTACACGATGAACCTCGGTTTCACTCGCG 181  
Qy 391 CCGACACGCTCGAGAAGGTGGCCGCTCTCTCGCGTGTGTGGCGCTGGCTTCCGCGGCTCT 450  
Db 182 CCGACACGCTCGAGAAGGTGATGTTGTTGGCCATGCTCACGCGTGGAGCCACCTCAGC- 240  
Qy 451 CTTCCCGCGCGGCTCGGGCTCGACTGGAGCATCAGCTCTTCTCTCTCCACGCTCC 510  
Db 241 -----CGCGCGGGGAGCCTCGAGTGGACCATCAGCTCTTCTCTCTCCAGGCTGC 292  
Qy 511 CCAACACGCTCGTATGGGATCCCGCTGTGTGGAGGATGACGCGCGCTGTCTGGCGG 570  
Db 293 CCAACACGCTCGTATGGGATCCCTTTGCTCAAGGGCATGTACGGGGAGTTCT---CCG 349  
Qy 571 GCACGCTCATGGTCAGGTCGTCGCTCTCCAGTGCATCATCTGTACACGCTCATGTCT 630



CC of auxin transport proteins and to design and/or identify specific  
XX inhibitors of auxin transport proteins, potentially useful as herbicides  
SQ Sequence 2162 BP; 428 A; 694 C; 636 G; 404 T; 0 U; 0 Other;

Query Match 19.4%; Score 536.2; DB 3; Length 2162;  
Best Local Similarity 59.4%; Pred. No. 11e-76;  
Matches 1021; Conservative 0; Mismatches 658; Indels 40; Gaps 5;  
QY 163 CCGCGCCATGATCAACGGCTGGAGCTCTACACAGTGTGACGGCGGTGGTGCCTGT 222  
DB 91 CTGAGGAGATGATATCCGGGACGACTTCTACAGGTGATGGGGGTGGTGCCTGT 150  
QY 223 AGTGGCCATGAGCTGGGTGAGCTGCGTCCGTGCTGGTGGGCGATCTTACCGCGGACC 282  
DB 151 AGTGGCGATGTTCTTGGGTGAGCTGCGTGGTGGTGGGCGATCTTACCGCGGACC 210  
QY 283 AGTGTCTCCGGGATCAACCGCTTCTGGGCGCTCTTCCGCGTGGCGCTCTCTCTCTTCACT 342  
DB 211 AGTGTCTCCGGGATCAACCGCTTCTGGCGCATCTTGGCGGTGGCGCTCTCTCTTCACT 270  
QY 343 TCATCTCCACCAACGACCCCTTTCGCCATGAACTCTGCTTCTGCGCGCGGACAGCTGC 402  
DB 271 TCATCTCCACCAACGACCCCTTTCGCCATGAACTCTGCGCTTCTGCGCGGACAGCTGC 330  
QY 403 AGAAGTGGCCCTCTTCCGGCTGCTGGGCTGGGCTCGGCGGCTCTCTCTCCCGCGCG 462  
DB 331 AGAAGTGGCTCTTCTGGCGGCTGCGCGGTGGTGGCGCTCTCTCTCGCGGACCGGCG 390  
QY 463 CGCTCGGCTGAGCTGAGCATCAGCTCTTCTCCCTCTCCACGCTTCCCGGCGGCTCG 522  
DB 391 CGCGCGGCTGAGCTGCTCATCAGCTCTTCTCCCTCTCCACGCTTCCCGGCGGACAGCTCG 450  
QY 523 TCATGGGATCCGCTGCTGCGAGCATGTAGCGGCGCTGCTGCGCGGCGCTCTCTCTG 582  
DB 451 TCATGGGATCCGCTGCTGATGCGCATGACG---GCCATCTCCGGCTGCTCTATGG 507  
QY 583 TCAGGTCTGCTCTTCCAGTCATCTGTATACAGCTCATGCTCTCTCTCTCTCTGAGT 642  
DB 508 TCAGATCTGCTGCTTCCAGTCATCTGTATACAGCTCATGCTCTCTCTCTCTGAGT 567  
QY 643 ACCGCGCGCGGCGCTGCTCTGAGCATGTTCCCGAGCGGCGCGCGGCTGCTCTCTG 702  
DB 568 TCCGCGCGCGGAGTGTGATCGCGACCACTTCCCGA---CACGGCGGCGTCCATCG 624  
QY 703 TCTCTCTCGCTCGACTCCGACGCTGCTCTGCTCGCGAGGGGAGCTCGAGCTCGAGG 762  
DB 625 TGTCTCTGACGTCGACCGCGAGCTGCTGCTGAGGGGCGGCCACCGGAGACGAGG 684  
QY 763 CCGAGCCCGAGCGGCTGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 822  
DB 685 CCGAGGTGGCGGACGCGGCGCTGCACTCACGCTGCGCGCTCTCTCTCTCTCTCTCT 744  
QY 823 GGTGCGGCTCACGCTGCGCAAGTCCACAGCTCGGCTCGAGGCGCGGCTCGCACT 882  
DB 745 GGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 804  
QY 883 CGCACTCCCAACCATGAGCCCGCTGTGTCCAACTCTTCCGCGGTGAGATCTACTCGC 942  
DB 805 GCTCTGCGGAAACCAACCCCGCGGCTTCACTTCAACCAACCGGCTCTCTCGGCA 864  
QY 943 TGCAGTCTGCGCAACCCCAACCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1002  
DB 865 TGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 924  
QY 1003 ACATCTGCGGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1062  
DB 925 CCTTCGAGCTTTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984  
QY 1063 CATGCGCGCGGCGGCGGAGACACTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCG 1122  
DB 985 ACAGACACTCGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1044

QY 1123 AGAGGAGGACCTGCACATGCTGCTGGAGCTCCAGCGCTCGCGCTGCTCCGAGCGG 1182  
DB 1045 ATGCCAAGGAGCTCCACATGTTCTGTTGAGCTCGAGCGCTCTCCGCTCTCAGAAGTCA 1104  
QY 1183 CCGCGCTGACGCTCTTC---GGCGCGGCGGCGCTGACCATGCGCGACGCTCTCGCAAAG 1239  
DB 1105 GCGCGCTGCTGCTGTTGAGTGGCGGCGGCGGCGGCGCTCTCGACGCTCGGCGCAAAG 1164  
QY 1240 GAGCCAGGCGCTACGAC-----GAGTACGGGCGG 1269  
DB 1165 AATCCACATGTTGATCCCGCGGACCTGCGCGAGAACCAACGGCTCAGGCAAGAGACG 1224  
QY 1270 ACGACTACAGCAGCAGGACGAGAACGGAGCGGCGGCGGAGCAAGGGCGGCGGCGG 1329  
DB 1225 AGAGTACCGCGCAGTGGCATTTGGGTGGCGGCGGCGGCGGAGAGAACTTCAGCTTCGAG 1284  
QY 1330 TGTGGAAGCTGGGGTCCAACTCGACGCGGCGAGCTGTACCCCAAGGACGAGCGGAGGGGA 1389  
DB 1285 GCGGCAAGAGCTGGAGCGGCGGCGGAGCAGTAGACGAGGAGCGGCTTGCCTGACGGG 1344  
QY 1390 GGGCGGCGGCTGGCGATGCGCGCGGAGCGGTGATGAGCGGCTCATCTCTCATG 1449  
DB 1345 TGACGAAGATGGGGTGGAGCTCGACGCGGAGCTGACCGGAGGTCGTCGACGTCGAGC 1404  
QY 1450 TGTGGAAGAGCTGATCCGGAACCCCAACACCTTACTCCAGCTCTATCGGCGCTCTGCT 1509  
DB 1405 GACCGAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1464  
QY 1510 CCTGCTCTCTTACAGGTGGGGCATCGAGATGCGAGCATCATGCGCGGCTCGATTTCGA 1569  
DB 1465 TGATGACACGCGCTCATCTCTCATTAATGGTGGCGAAGCTCATCCGCAACCCCAACT 1524  
QY 1570 TCTGTGCGGACGCGGCTGCGGATGCCATGTTTCAAGCT-AGGCTGTTCATGGCGCTG 1628  
DB 1525 ACTCGAGCTCTCGGCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1584  
QY 1629 CAGCGAGGATCATCGCTGCGGGAACAAGCTGGCGGCATCGCGATGGGCGCTCGGCTTC 1688  
DB 1585 CAGCCAGCATCATCGCTGCTGGCAATCAGCGCGCTGCTCTCCATGGCGGCTCGGCTTC 1644  
QY 1689 GTGCGAGGCGCGGCTGATGGCGCGCTCCATCGGCTCGGCTGCTGCGCGGCTCGCTC 1748  
DB 1645 CTGCGGCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1704  
QY 1749 CTCACATCCCATCGTCCAGGCTGCTCTCTCTCAGGGAATCGTCCGTTGCTGCTTCGC 1808  
DB 1705 CTGACGTCGCCATTTGTCAGGCGCTTACCAAGGAGTGTGCTTTGTTTGTGA 1764  
QY 1809 AAGAGTACGGCTTTCATCCGACATCCCTGAGCAGAG 1847  
DB 1765 AAGAATACAAATGTCCACCGGCGCATCTGAGCAGAG 1803

## RESULT 10

AAA94720

ID AAA94720 standard; DNA; 504 BP.

XX AAA94720;

AC AAA94720;

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX 02-FEB-2001 (first entry)

XX	07-MAY-1999;	99US-0133040P.	
XX	(DUPO ) DU PONT DE NEMOURS & CO E I.		
XX	(PION-) PIONEER HI-BRED INT INC.		
PA			
PA			
PI	Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;		
PI			
XX	WPI; 2000-687647/67.		
XX	P-PSDB; AAB26933.		
XX			
PT	New nucleic acid sequences encoding new auxin transport proteins, useful		
PT	for modulating root growth of plants and to screen for herbicides.		
PT			
XX	Claim 2; Page 59; 94pp; English.		
XX			
CC	Auxins are plant hormones that influence plant behaviour and development		
CC	e.g. vascular tissue differentiation, apical development, tropic		
CC	responses and organ (e.g. flower, leaf) development. The present		
CC	invention relates to corn auxin transport protein coding sequences and		
CC	proteins. The present sequence is one such coding sequence. This sequence		
CC	may be used to modulate root development, e.g. to produce a more robust		
CC	root system, alter root angle or redirect root growth. Also, the present		
CC	sequence may be useful for gene mapping (e.g. for plant breeding) and to		
CC	identify loss of function mutants. The protein encoded by the present		
CC	sequence may be useful for raising specific antibodies, for the detection		
CC	of auxin transport proteins and to design and/or identify specific		
CC	inhibitors of auxin transport proteins, potentially useful as herbicides		
XX			
XX	Sequence 504 BP; 68 A; 202 C; 110 G; 108 T; 0 U; 16 Other;		
XX			
XX	Query Match 16.3%; Score 452.6; DB 3; Length 504;		
XX	Best Local Similarity 95.6%; Pred. No. 2.3e-63;		
XX	Matches 477; Conservative 0; Mismatches 20; Indels 2; Gaps 2		
Qy	13 TGAGCCCTACAAACACTCTCTTCTTTCATTTGCTCCACACTACCATCTCATCTCTCGGCCAT 72		
Db	6 TGAGCCCTACAAACACTCTCTTCTTTCATTTGCTCCACACTACCATCTCATCTCTCGGCCAT 65		
Qy	73 TTTTACACCACTCCCTTCTCGTTGTGCAACCCAAACAAATTGGCACTGCTGCGCGACCCCT 132		
Db	66 TTTTACACCACTCCCTTCTCGTTGTGCAACCCAAACAAATTGGCACTGCTGCGCGACCCCT 125		
Qy	133 CTTCTCCTCCCGCGTCTCCCGACAAAGCCATCCGCGGCCATGATCAACGGCTTGACCTCT 192		
Db	126 NCTTCTCCTCCCGCGTCTCCCGACAAAGCCATCCGCGGCCATGATCAACGGCTTGACCTCT 185		
Qy	193 ACCAGTGTGTACGGC -GGTGGTCCGCTGTAGTGGCCATGACGCTGCGGTACGGCTCC 251		
Db	186 ACCACNGCTGTACGGCTGGNGGTCCGCTGTACTGGGCCATGACGCTGGGACGNGTNC 245		
Qy	252 GTCCGCTGTGTGGCGCATCTTTCACGCCGAGACAGTGTCTCCGGGATCAACCGCTTCGTGGCG 311		
Db	246 GTCCGCTGGNGGNGCATWNTNCAACCGGACAGTGTCTCCGGNATCAACCGCTTCGTGGCG 305		
Qy	312 CTCTTTGGCGGTGCGCTCTCTCTCTTCCACTTTCATCTCCACCAACACCCCTTCGCCATG 371		
Db	306 CTCTTTGGCGGTGCGCTCTCTCTCTTCCACTTTCATCTCCACCAACACCCCTTCGCCATG 365		
Qy	372 AACCTGCGCTTCTTGGCGCGGACACGCTGCAGAAAGTGGCCGCTCTCTCGCGTGTGTGGCG 431		
Db	366 AACCTGCGCTTCTTGGCGGTGCACACGCTGCAGAAAGTGGCCGCTCTCTCGCGTGTGTGGCG 425		
Qy	432 CTGGGCTTCCCGGGGCTCTCTCTCCCGGGCGCGTCC -GGGCTCGACTGAGGATCACGCT 490		
Db	426 CTGGGCTTCCCGGGGCTCTCTCTCTCCCGAGNGCGCTCAGGGCTCGACTGAGGATCAAAGCT 485		
Qy	491 CTCTTCCCTCTCTCCAGCTC 509		
Db	486 CTTCNCCCTCTTCCAGCTC 504		
XX			

AA94719	ID	AAA94719	standard; cDNA; 1426 BP.
XX	XX	AA94719;	
XX	AC	AA94719;	
XX	DT	02-FEB-2001	(first entry)
XX	XX		Corn auxin transport protein clone p0094.csssh17r DNA sequence.
XX	DE		Auxin transport protein; corn; root development; gene mapping;
XX	KW		plant breeding; herbicide; ss.
XX	KW		Zea mays.
XX	OS		
XX	XX		
XX	EH	Key	Location/Qualifiers
XX	FT	CDS	320..1426
XX	FT		/*tag= a
XX	FT		/product= "Corn auxin transport protein clone
XX	FT		p0094.csssh17r"
XX	FT		/note= "No stop codon given"
XX	XX		
XX	PN	W0200068389-A2;	
XX	XX		
XX	PD	16-NOV-2000.	
XX	XX		
XX	PF	03-MAY-2000;	2000MO-US012061.
XX	XX		
XX	PR	07-MAY-1999;	99US-0133040P.
XX	XX		
XX	FA	(DUPO ) DU FONT DE NEMOURS & CO E I.	
XX	FA	(PION-) PIONEER HI-BRED INT INC.	
XX	XX		
XX	PI	Orozco EM, Weng Z, Bruce WB, Cahoon RE, Tao Y;	
XX	XX		
XX	DR	WPI; 2000-687647/67.	
XX	DR	P-PSDB; AAB26932.	
XX	XX		
XX	FT	New nucleic acid sequences encoding new auxin transport proteins, useful	
XX	FT	for modulating root growth of plants and to screen for herbicides.	
XX	XX		
XX	PS	Claim 2; Page 56; 94pp; English.	
XX	XX		
XX	CC	Auxins are plant hormones that influence plant behaviour and development	
XX	CC	e.g. vascular tissue differentiation, apical development, tropic	
XX	CC	responses and organ (e.g. flower, leaf) development. The present sequence	
XX	CC	is a cDNA insert comprising of clone p0094.csssh17r. This sequence	
XX	CC	encodes a corn auxin transport protein. This sequence may be used to	
XX	CC	modulate root development, e.g. to produce a more robust root system,	
XX	CC	alter root angle or redirect root growth. Also, the present sequence may	
XX	CC	be useful for gene mapping (e.g. for plant breeding) and to identify loss	
XX	CC	of function mutants. The protein encoded by the present sequence may be	
XX	CC	useful for raising specific antibodies, for the detection of auxin	
XX	CC	transport proteins and to design and/or identify specific inhibitors of	
XX	CC	auxin transport proteins, potentially useful as herbicides	
XX	XX		
XX	SQ	Sequence 1426 BP; 241 A; 528 C; 415 G; 242 T; 0 U; 0 Other;	
		Query Match	15.4%; Score 425.6; DB 3; Length 1426;
		Best Local Similarity	72.1%; Pred. No. 5.2e-59;
		Matches 646; Conservative	0; Mismatches 199; Indels 51; Gaps 5;
Qy	171	ATGATACACCGGTGGACCTTACACGCTGCTACCGCGGTGTGCGCTGTACGTGGCC	230
Db	320	ATGATTACGGGGACGAGCTTCTACACGTCATACCGCGGTGTGCGCTGTACGTGGCG	379
Qy	231	ATGACGCTGGCGTACGCTCCGCTCGGTGGCGGATCTTACGCGCGACCAAGTGTCC	290
Db	380	ATGATCTGGGCTACGGGTGGTGGCGGATCTTCTCGCGGACCAAGTGTCTCC	439
Qy	291	GGGATCAACCGCTTGTGGCGCTCTTCGCGCGTGGCGCTCTCTCTCTTCATCTCC	350
Db	440	GGGATCAACCGCTTGTGGCGCTCTTCGCGCGTGGCGCTCTCTCTTCATCTCC	499

RESULT 11







131 CTCCTCCCTCCCGCGTCCCGGCAAGCAATCCGCGCCATGATCACCGCGCTGGACCT 190  
Db 62 CTCCTCTCTCTCTCTCTCTCTCTCCGCGCAACAAAGAGATGATACGGCGGCGACTT 121  
Qy 191 CTACCACTGCTGACGGCGGTGGTCCCGCTGTACGTGCCATGACGCTGGGCTACGGCTC 250  
Db 122 CTACCACTGATGACGGCTATGGTTCCTGTTATACGTAGCTATGATCCCTGCTTACGGCTC 181  
Qy 251 CGTCCGCTGGTGGCGCATCTTTACGCCGACCAAGTGTCTCCGGGATCAACCGCTTCGTGGC 310  
Db 182 TGTCAATGGTGGAAATCTTACACAGACCAATGCTCCGGCATAAACGCTTTCGTGCG 241  
Qy 311 GCTCTCCGCGGTGCGCTCTCTCTCTCCATCTTCACTTCAACAGACCCCTTCGCGAT 370  
Db 242 TCTCTTCGCGTTCCT 301  
Qy 371 GAACTCGCTTCCTGGCGCGGACACGCTGACAGAGGTGCGCTCTCCGCGTCTGCTGGC 430  
Db 302 GAACTCGCTTCCTGGCGCGGATCTCTCTCAAGAACTCATTTCTCTCTCTCTCTCTCT 358  
Qy 431 GCTGCGCTCCGCGGCT 490  
Db 359 -----CTTCTCTCTGTAACACTCAGCGCGCAACGGTTCTTTAGATTGACCATAACTCT 412  
Qy 491 CTTCTCCCTCTCCAGCTCCCGCAACAGCTGCTGATGGCATCCGCTGCTCGAGGCAT 550  
Db 413 CTTCTCTCTCTCGACACTCCCGCAACACTCTAGTATGGGATACCTCTCTCTCAAGGCAT 472  
Qy 551 GTACGCGCGCTCGTGGCGCGGACGCTCATGTCTCAGGTCTGCTCTCCAGTGCATCAT 610  
Db 473 GTATGGTAATTCT--COGGGACCTCATGTTCAAACTGTTGTTCTTCAAGTGTATCAT 529  
Qy 611 CTGGTACAGCTCATGCT 670  
Db 530 TTGGTACATACATCATGCT 589  
Qy 671 CCAGTTTCCCGACGGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 730  
Db 590 GCAGTTTCCAGACA--CAGCAGGATCTATTGTTTCGATTATGTTGATTTCGACATTTAT 646  
Qy 731 CTCGCTCGCAGGGGGAGC---TCGAGCTCGAGGCGCGAGCGCGC-----777  
Db 647 GTCTTTAGATGGAAGACAACCTTTGGAACTGAAGCTGAGATTAAAGAAATGGGAAGCT 706  
Qy 778 -----TCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 818  
Db 707 TCATGTTACTGTTGCTGTTCTTAATGCTTCAAGGCTCTGATATTTACTCGAAGTCTCA 766  
Qy 819 GGGCGGGTCCGCTCACCGTGGCAAGTCCACAGCTCGCGCTCCGAG-----GCC 869  
Db 767 AGGCTTATCTCGACACCTTAGACCTTCGAATCTAACCAACGCTGAGATATATTCGCTTCA 826  
Qy 870 GGGTGTCTGCATCTGGACTCCAGACCATGAGACCCCGGTGTCTCAACCTCTCCGGCGTG 929  
Db 827 GAGTTTCAAGAAACCCCAACCCCAACCGTGGCTCTAGTTTAAATCATACTGATTTTACTCGAT 886  
Qy 930 GAGATCTACTGCTGACGCTGTCGGCGCA-----958  
Db 887 GATGGCTTCTGGTGGTGGTGGAACTCTAACTTTGCTCTGGAGAAGCTGTGTTGGTTC 946  
Qy 959 -----CCCCACCCCGCGGGTTCAGCTTCAACCAACCGCGACTTCTTCAACATCGTCTGG 1012  
Db 947 TAAAGTCTCTACTCCGAGACCTTCCAACTACGAAGAAGACGCTGCTCTCTTAAACCGAC 1006  
Qy 1013 CGCGCGCGCAAGGAGGCGGAGGAGCGCGGGGACAGGAGAGGCGCATCGCGGG 1072  
Db 1007 GGTGCTGGAATGCTGCTGCTGAGCTGGAGGTTTCAATATCAATCTGAGGAGTGGTGG 1066  
Qy 1073 CGCGCGGAGGACACTCCGCGGACCGCGAGCGCGTCTGCGCGGCAAGA-----1125  
Db 1067 CGGTGAGGAGCGCAATTATCCGCGCGCGAACCAGGAGTGTCTTCCGCCAACACTGGCGG 1126

Qy 1126 -----1125  
Db 1127 TGGTGGAGCACGGCGCGAAAGAAACGCTCCGGTGGTTGGTGGAAAGACAAGACGG 1186  
Qy 1126 -----GGAAGGACCTGCATGCTGCTGGAGCTCCAGCGCTCGCCGCTGTCCGAGCG 1180  
Db 1187 AAACGGAAGAGATCTTTCATGTTTGTGTGGAGCTCAAGTGTTCGCCCGGTCTCAGA---1243  
Qy 1181 CGCGCGCGTGCACGCTTTCGGCGCGCGCGCTGACCAATGCGGAGCTCTCGCCAAAG 1240  
Db 1244 -----TGTGTTCCGCGGTGGAGGAGAAACCAACACGCGGATTAATCCACCGCTACGAA 1297  
Qy 1241 AGCCAGGCGCTACGACGAGTACGGGCGCGACGACTACACAGCAGGACCAAGAAACGGAG 1300  
Db 1298 CGATCATCAAAAGAGCTTAAAGTCTCTGTACCTCAGGGAAATAGTAAAGCAACACAGTA 1357  
Qy 1301 CGCGCGCGGGAACAAGGGCGGCGCGCTGCTGCAAGCTGGGGTCCAACTCGACGCGCA 1360  
Db 1358 CGTGGAGAGGAAGAGTTTAGTTTCGGTAAACAAAGACGATGATAGCAAAAGTATGGCAAC 1417  
Qy 1361 GCTGTACCCAGGACGACGGCGAGGGGAGGGCGGCGGTGGCGATCGCGCGCGGAG 1420  
Db 1418 GACGCTGGGAAACAACATAAGCAACAAACGACGCGAGGTAAAGTGTATGCAACAAAG 1477  
Qy 1421 CGTGATGACGCGGCTCATCTCTCATGCTGTGGAGAAAGCTGATCCGGAACCCCAACAC 1480  
Db 1478 TGTGATGACAAGACTCATTTCTCATTTAGTTTGGAGAAACTTATTCGTAATCCCACTC 1537  
Qy 1481 CTACTCCAGCTCATCGGCTGCTGTGCTCTCTGCTCTCTACAGTGGGCAATCGAGAT 1540  
Db 1538 TTACTCCAGTTTATTCGGCATCACCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1597  
Qy 1541 GCCAGGATCATCGCGCGCTCGATTTTCGATCTCTGCGACGCGGCTCTCGGATGGCCAT 1600  
Db 1598 GCCAGCTCTTATAGCAAGTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1657  
Qy 1601 GTTCAGCTAGGCTGTTCATGCGCTGCGAGCGAGGATCACTGCGTGGGGAACAGCT 1660  
Db 1658 GTTCAGTCTTGGGTTGTTTCATGCGCTTAAACCCCAAGATAATAGCTTGTGGAAACAGA 1717  
Qy 1661 GCGGCGCATCGGATGGGCTCGGTTCTGTCGAGCGCGCGGCTCATGCGCGCGCTC 1720  
Db 1718 AGCAGCTTTTTCGGCGGCTATGAGATTGTCGTTGAGCTGCGCTCATGCTCTGTTGCTTC 1777  
Qy 1721 CATCGCGCTCGTCTGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1780  
Db 1778 TTATGCGTTGGCTTCGTTGGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1837  
Qy 1781 TCAGGAGATCGTCCGTTGTTGTTGTCGCAAGGATGACGGGTTTCATCCGACATCTGAG 1840  
Db 1838 GCAAGGAATAGTACCGTTTGTGTTGCCAAAGATATAATGTGCTATCTCTGACATCTTAG 1897  
Qy 1841 CACAGCG 1847  
Db 1898 CACTGCG 1904

## RESULT 14

AAA94716

ID AAA94716 standard; cDNA; 1088 BP.

XX AAA94716;

AC AAA94716;

XX

DT 02-FEB-2001 (first entry)

XX Corn auxin transport protein clone contig #1 DNA sequence.

XX Auxin transport protein; corn; root development; gene mapping;

XX plant breeding; herbicide; ss.

XX Zea mays.

OS

XX

FH

Key Location/Qualifiers



QY 351 ACCAACGCCCTTCGCGATGAACCTCGCTTCCTGGCGCGGACACGCTCGAGAAGTG 410  
Db 181 TCRAACAATCTTACGAAATGAATCTCGAATTCATCGCAGCGGATACACTCCAAATACTA 240  
QY 411 GCCGCTCGCGCTCTCGCGCTCGCGCTCCCGCGCTCTCTCTCCCGCGCGCTCGGG 470  
Db 241 ATCATGCTCACTCTCTCT-----AATCATCTGGGCTAACTTCACTCGCTCGCGTAGT 291  
QY 471 CTCGACTCGAGCATACAGCTCTTCTCCCTCTCCCTCCAGCTCCCAACAGCTCGTATGGG 530  
Db 292 CTTGAATGAGCATACAAATCTTCTCTCTTCAACACTTCCCAACACTCTCGTTATGGGA 351  
QY 531 ATCCGCTCTCTCGAGGATGACGCGCTGTCGCGCGGACGCTCATGCTCGAGGTC 590  
Db 352 ATACCTCTCTTATCGCCATGACGGAATATT---CGGCTCACTCATGCTTCAGATC 408  
QY 591 GTCGCTCTCCAGTGCATCATCTGGTACAGCTCATGCTCTCTCTCTCTCGAGTACGCGCC 650  
Db 409 GTTGCTCTTCACTGATATATCTGGTACACGCTCTCTCTCTCTCTCTCTCTCGAATACAGGA 468  
QY 651 GCGCGCGGCTCGTCTCTCGACAGTTCGCGAGCGCGCGGCTCCATGCTCTCTCTTC 710  
Db 469 GCTAAGATCTTGATCATAGGAAGTTTCAGA---GACTGGTCTTCGATTTGATCTGTC 525  
QY 711 CCGCTCGACTCGAGCTGCTCTCGCTCGCGAGGGGAGCTCGAGCTCGAGCGCGAGCCC 770  
Db 526 AAAAGTTGATCGATGTTGTTCTTTTAGATGACATGATTTTCTTGAACACTGATGCTCAA 585  
QY 771 GACGCGCTCGCGCGCGCGCGCTCTCTCTCCCGCGCGGAGCGCGCGGCTGCGG 830  
Db 586 ATCGGTGACGATGTTAGCTTCACTGTTAGCTGAGAAATCAACGCTTCTC-GGAGATC 644  
QY 831 GTACCGTGGCAATGCCACAGCTCGGCTCCGAGCGCGGCTGCTCGCACTCG----- 884  
Db 645 TTTTACGCTGGTGGTGTACTAAATGATGACTCTCGTCTGCTAATCTCACCGGAGTGA 704  
QY 885 -----CACTCCAGACCATGACG----- 903  
Db 705 GATTTATAGCTTAACACTACTTCAAGAGAGTAACTTCAATCAATCTGATTTTATCTC 764  
QY 904 -----CCGCTGTGTCCAACCTCTCCGCGCTGGAGATCTACTCGCT 943  
Db 765 TATGATGGGGTTTCCGCGTGGCGGCTTTGGAACCTTTGCTCGCGGATATGTACTCTGT 824  
QY 944 GAGTCTGTCGCGCAACCCACCGCGCGGCTCGAGCTTCAACACGCGGACTTCTTCAA 1003  
Db 825 TCAGTCTCTCGAGGAGCTACTCTCTAGACCTTTGGAATTTTGAGGA-GAGTTGCTATGG 883  
QY 1004 CATGCTCGCGCGCGCGCAAGGAGCGGAGAGCGCGGCGGACGAGGAGAGGCGC 1063  
Db 884 CTTCTTCGCAAGATTCGGGTATTACCGGAGGAGCTCCCGGCTCGTACCCAGCTCTTA 943  
QY 1064 ATCGCGCGCGCGCGGAGGACACTCGCGCGAGCGCGCGAGCGCTCGCGCGCGC--- 1120  
Db 944 ATCCGAGTTTTTCTACCGGTAAATAAACCAGTAGTAAAGCTCCAAAGAGAAATCATATC 1003  
QY 1121 -----CAAGAGGAGACTTCACATGCTCTCTGGAGCTCCA 1158  
Db 1004 ATGTAGGAAATCAATAGTAATGATGCTAAGAGCTTCAATGTTTGTGGGATCCA 1063  
QY 1159 GCGCTCTCGCCCTGTCGAGCGCGCGCGCTGTCACGT-----CTTCGCGC 1203  
Db 1064 ACGGATCACCCGTTTCGGACCGAGCTGCTTCAAGTTGATAATGGAGCAATGAACAAG 1123  
QY 1204 CCGCGCGGCTGACCATGCGAGCTCTCTCGCCAAAGAGCCCGCTACGACGAGTACG 1263  
Db 1124 TCGGAAATCCGATCAAGCGCTGCAAAAGAGATTGGAATGTTGATCTCTGATCATATCTC 1183  
QY 1264 GCGCGGACCACTTACAGCAGCAGAGAGAAACGAGGAGCGCGCGGACAGGCGCGGCG 1323  
Db 1184 AAAATGGTGAACACAAAGCTGCTCCGATGAACGCGGACTATGGCGGGGAGAGAGTCGG 1243  
QY 1324 CGAG-----CTGTCGAAGCTGGGCTCCAACTCGAGCGGCGCAGC 1362

Db 1244 AGAGGTAAGGAAGTGCTTAACGGACTACACAGCTTCGGTGTAACTCCACAGCAGAGC 1303  
QY 1363 TGTACCCCAAGGACGAGCGGAGGAGCGCGCGCGGTGG-----CGATGCCGCGCG 1416  
Db 1304 TAAACCCCTAAGGAAGCTATAGAAAACGGGTGAAACTGTACCGGTAAACATATATGCCACAG 1363  
QY 1417 CGAGCGTGTATGACCGCGCTCATCTCATCATGTGTGTGAGGAAGCTGATCCGGAACCCCA 1476  
Db 1364 CGAGTGTGATGACTCGCTGATATTGATTAATGGTGTGAGGAACCTCATAGAAGAACCCAA 1423  
QY 1477 ACACCTACTCCAGCTCATCGCGGCTGTGTGTCCTGTCTCTTACAGGTGGGGCATCG 1536  
Db 1424 ACACCTACTCTAGTCTCATTTGGCTCTCATTTGGCTCTTGTGCTTTTCAGGTGGGATGTGG 1483  
QY 1537 AGATGCCAGCATCATCGCGCTCGATTTTCGATCTGTGAGACGGGCTCTCGGATGG 1596  
Db 1484 CAATGCTTAAATATTCAACATCAATCTCAATTTCTTCTGATGCTGTCTTGTGTATGG 1543  
QY 1597 CCATGTTCAAGCTTAGGCTGTTCATGGCTGTGAGCGGAGATCATCGGCTGGGGAACA 1656  
Db 1544 CAATGTTCAAGTTTGGGTTGTTTCATGGCATTTCAACCGAAATTAATTTGCTTGGGAAAT 1603  
QY 1657 AGTGGGCGCATCGCATGGGCTCGGTTCTGTCGAGCGCGGCTCATGGCGCGCG 1716  
Db 1604 CGACGCGACTTTTGGATGGCGGTGAGATCTTTACTGACACGCGGTAAATGGCCGTGG 1663  
QY 1717 CTTCCATCGCGCTCGGCTCTCGCGCGCTCTCTCTCCACATCGCCATCGTCCAGGCTGCTC 1776  
Db 1664 CAGCAATGGCTATTGGATTACGTGGAGACCTATTGGCTGTGGCCATTGTTCAAGCTGCAT 1723  
QY 1777 TGCTCAGGGATCTGTCGCTTCTGTTTCCGCAAGAGTACGCGCTTCATCCGACATCC 1836  
Db 1724 TGCTCAGGGATCTGTCGCGCTTTGTTTTCGAAAAGAGTATAATGTTTCATCCCGCAATCT 1783  
QY 1837 TGAGCACAGCG 1847  
Db 1784 TGAGTACAGG 1794

Search completed: March 3, 2004, 01:25:04  
Job time : 1062 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: March 3, 2004, 01:10:58 ; Search time 6898 Seconds  
(without alignments)  
11987.297 Million cell updates/sec

Title: US-10-030-884-13

Perfect score: 2769

Sequence: 1 ccacgcgtccgctgagccct.....aaaaaaaaaaaaaaaaaaag 2769

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba.\*

2: em\_esthum.\*

3: em\_estin.\*

4: em\_estmu.\*

5: em\_estov.\*

6: em\_estpl.\*

7: em\_estro.\*

8: em\_hic.\*

9: gb\_est1.\*

10: gb\_est2.\*

11: gb\_hic.\*

12: gb\_est3.\*

13: gb\_est4.\*

14: gb\_est5.\*

15: em\_estfun.\*

16: em\_estom.\*

17: em\_gss\_hum.\*

18: em\_gss\_inv.\*

19: em\_gss\_pln.\*

20: em\_gss\_vrt.\*

21: em\_gss\_fun.\*

22: em\_gss\_mam.\*

23: em\_gss\_mus.\*

24: em\_gss\_pro.\*

25: em\_gss\_rod.\*

26: em\_gss\_phg.\*

27: em\_gss\_vrl.\*

28: gb\_gss1.\*

29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2108	76.1	2737	11	AY110494
2	877.8	31.7	893	29	CG376782
c	3	840	30.3	29	CG358070
4	726	26.2	726	29	CG358058

C	5	711	25.7	730	29	CG376770	CG376770
	6	706	25.5	717	14	CF244269	CF244269
	7	639	23.1	835	29	CC722061	CC722061
	8	619.4	22.4	658	13	BU098540	BU098540
c	9	599	21.6	958	28	CC371169	CC371169
	10	576.8	20.8	610	13	BU080190	BU080190
	11	574.4	20.7	630	14	CA830783	CA830783
	12	569	20.5	627	13	BU037419	BU037419
	13	568.4	20.5	794	29	CG456385	CG456385
	14	561	20.3	776	29	CG317974	CG317974
	15	559.8	20.2	604	14	CA831110	CA831110
	16	527.4	19.0	815	28	BZ644423	BZ644423
	17	518.4	18.7	552	13	BU037594	BU037594
	18	510	18.4	510	14	CF244718	CF244718
	19	477.6	17.2	497	13	BU036965	BU036965
	20	431	15.6	469	14	CD001628	CD001628
	21	429	15.5	857	29	CG057348	CG057348
	22	429	15.5	891	29	CG057347	CG057347
	23	425.6	15.4	1426	11	AY106004	AY106004
	24	413.8	14.9	1107	14	CK209475	CK209475
	25	399	14.4	831	28	BZ644430	BZ644430
	26	396.2	14.3	903	14	CF243672	CF243672
	27	395.4	14.3	1105	14	CK208792	CK208792
	28	385.4	13.9	1109	14	CK208849	CK208849
	29	383.4	13.8	428	13	BU500094	BU500094
	30	382.8	13.8	386	28	BH801495	BH801495
	31	380.4	13.7	1221	11	AY110024	AY110024
	32	380.2	13.7	924	29	CG293496	CG293496
	33	378.6	13.7	722	14	CD934785	CD934785
	34	376.4	13.6	619	14	CA326508	CA326508
	35	364.6	13.2	533	28	BH873074	BH873074
	36	359.8	13.0	696	14	CA184766	CA184766
	37	356.4	12.9	739	14	CB645882	CB645882
	38	356.4	12.9	819	14	CB648277	CB648277
	39	350.4	12.7	703	29	CG637427	CG637427
	40	349.8	12.6	899	14	CA227062	CA227062
	41	346.2	12.5	585	13	CA144854	CA144854
	42	344.6	12.4	731	14	CB668851	CB668851
	43	342.2	12.4	631	13	BQ295136	BQ295136
	44	339.6	12.3	621	13	CA099711	CA099711
	45	338.8	12.2	694	14	CA192422	CA192422

#### ALIGNMENTS

RESULT 1	AY110494	2737 bp	mrna	linear	HTC 17-OCT-2002
LOCUS	AY110494	2737 bp	mrna	linear	HTC 17-OCT-2002
DEFINITION	Zea mays CL464_1 mRNA sequence.				
ACCESSION	AY110494				
VERSION	AY110494.1	GI:21214903			
KEYWORDS	HTC.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.				
REFERENCE	1 (bases 1 to 2737)				
AUTHORS	Rainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whittitt, M.S., Arthur, L.W., Hanafey, M., Morgante, M. and Tingey, S.V.				
TITLE	Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes				
JOURNAL	Unpublished (2002)				
REFERENCE	2 (bases 1 to 2737)				
AUTHORS	Coe, E.H.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA				
COMMENT	If you are interested in getting corresponding physical clones, these are publicly available from ZmDB and may be found by BLAST searching at MSL, <a href="http://maizemap.org">maizemap.org</a> ; ZmDB, <a href="http://www.zmdb.iastate.edu">www.zmdb.iastate.edu</a> ; TIGR, <a href="http://www.tigr.org">www.tigr.org</a> ; or NCBI, <a href="http://www.ncbi.nlm.nih.gov">www.ncbi.nlm.nih.gov</a> . When the source of the				



```
QY 1861 CATCGCATGTTTTCATCACTGGCCATAGTTAAACGGGAAAAAAGCAGACGATCGATG 1920
Db 1861 CATCGCATGTTTTCATCACTGGCCATAGTTAAACGGGAAAAAAGCAGACGATG 1920
QY 1921 ACAGCGCACTGAATTCATCACTGATTAATTAATGATGTTTTCATGATGATG 1980
Db 1921 ACAGCGCACTGAATTCATCACTGATTAATTAATGATGTTTTCATGATGATG 1980
QY 1981 AAGAACCACTAATTAAGCACTGATCTAGGACAGATCAGATGATGATGATGTTTC 2040
Db 1981 AAGAACCACTAATTAAGCACTGATCTAGGACAGATCAGATGATGATGATGTTTC 2040
QY 2041 TCCTGACAACTGATTTTCTTACTACACAGTGTGCTTCACTATCCATCCAGATCA 2100
Db 2041 TCCTGACAACTGATTTTCTTACTACACAGTGTGCTTCACTATCCATCCAGATCA 2100
QY 2101 TACAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
Db 2101 TACAACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
QY 2161 CATCGCTCTGCGCCATCACTGCTACTATCTTCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db 2161 CATCGCTCTGCGCCATCACTGCTACTATCTTCTGCTGCTGCTGCTGCTGCTGCTG 2220
QY 2221 TCGCTTCTCAGCTGCGGAGCGCCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
Db 2221 TCGCTTCTCAGCTGCGGAGCGCCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2280
QY 2281 TGATGACATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
Db 2281 TGATGACATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
QY 2341 CGGACCTGGGGAGTTCAGACAGACACCCCTCGAACAAGATTTCTTTTGGCTTCTGTC 2400
Db 2341 CGGACCTGGGGAGTTCAGACAGACACCCCTCGAACAAGATTTCTTTTGGCTTCTGTC 2400
QY 2401 GTGAGAACAAAGTTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Db 2401 GTGAGAACAAAGTTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
QY 2461 ATCATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
Db 2461 NNNTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
QY 2521 CCAGAGCAATGGAATCAACAAAGCTTCTGCGAGCGGTATCATATCAAGGAACACA 2580
Db 2521 CCAGAGCAATGGAATCAACAAAGCTTCTGCGAGCGGTATCATATCAAGGAACACA 2580
QY 2581 TGCAGATACAAACGAGTCTAGTGCAGATGCGCAATGGCTTCTTCTTTTCTTGGGAAA 2640
Db 2581 TGCAGATACAAACGAGTCTAGTGCAGATGCGCAATGGCTTCTTCTTTTCTTGGGAAA 2640
QY 2641 GGGTTTCTAGATGATTAAGGATTCAGATGATGATGATGATGATGATGATGATGATG 2700
Db 2641 GGGTTTCTAGATGATTAAGGATTCAGATGATGATGATGATGATGATGATGATGATG 2700
QY 2701 ACAATTTTCTGGCTTTTATGAAAATTCCTCTCGTT 2737
Db 2701 ACAATTTTCTGGCTTTTATGAAAATTCCTCTCGTT 2737

RESULT 2
CG376782
LOCUS CG376782 893 bp DNA linear GSS 26-AUG-2003
DEFINITION CG376782 Zea mays genomic clone ZMMBma0776C20,
genomic survey sequence.
ACCESSION CG376782
VERSION CG376782.1 GI:34294049
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
source
location/Qualifiers
1..893
/organism="Zea mays"
/mol_type="genomic DNA"
/strains="B73"
/db_xref="taxon:4577"
/clone="ZMMBma0776C20"
/clone_lib="ZM_0.7_1.5_KB"
/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb
methylation filtered genomic DNA library"

ORIGIN
Query Match 31.7%; Score 877.8; DB 29; Length 893;
Best Local Similarity 99.7%; Pred No. 9,6e-80;
Matches 890; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 103 ACAAAATTGGCACTGCTGTGCGGACCCCTCTCTCCCTCCCGCGTCCCGGACGAT 162
Db 1 ACAAAATTGGCACTGCTGTGCGGACCCCTCTCTCCCTCCCGCGTCCCGGACGAT 60

QY 163 CGCGGCGCATGATACCGCGTGGACCTTACACGCTGTACGCGGTGTGCGCGTGT 222
Db 61 CGCGGCGCATGATACCGCGTGGACCTTACACGCTGTACGCGGTGTGCGCGTGT 120

QY 223 AGCTGGCCATGACGCTGGCGTACGCGTCCGTCGCTGTGCGGATCTTCAAGCGGACC 282
Db 121 AGCTGGCCATGACGCTGGCGTACGCGTCCGTCGCTGTGCGGATCTTCAAGCGGACC 180

QY 283 AGTGTCTCGGATGATCAACCGCTTCTGCGGCTCTTCCGCTGCGCTCTCTCTTCCACT 342
Db 181 AGTGTCTCGGATGATCAACCGCTTCTGCGGCTCTTCCGCTGCGCTCTCTCTTCCACT 240

QY 343 TCATCTCCACCAACGACCCCTTCGCGCATGAACCTGCGCTTCTGCGCGCGGACGCTGC 402
Db 241 TCATCTCCACCAACGACCCCTTCGCGCATGAACCTGCGCTTCTGCGCGCGGACGCTGC 300

QY 403 AGAAGTGGCGCTCTCTCGGCTGTCTGGCGTGGCTTCCGCGGCTCTCTCTCCCGCGG 462
Db 301 AGAAGTGGCGCTCTCTCGGCTGTCTGGCGTGGCTTCCGCGGCTCTCTCTCCCGCGG 360

QY 463 CGCTCGGCTCGACTGAGCATCAAGCTTCTCTCTCTCCCTCTCCACGCTCCCGGACGCTCG 522
Db 361 CGCTCGGCTCGACTGAGCATCAAGCTTCTCTCTCTCCCTCTCCACGCTCCCGGACGCTCG 420

QY 523 TCATGGGCAATCCCGTGTGTGAGGATGTACGGCGGTGTGCGCGGCTGTGCGCGGCTCATGG 582
Db 421 TCATGGGCAATCCCGTGTGTGAGGATGTACGGCGGTGTGCGCGGCTGTGCGCGGCTCATGG 480

QY 583 TCCAGTGTCTCTCTCCAGTGCATCTGTGTACAGCTCATGTCTCTCTCTCTCTCTCTCTCT 642
Db 481 TCCAGTGTCTCTCTCCAGTGCATCTGTGTACAGCTCATGTCTCTCTCTCTCTCTCTCTCT 540

QY 643 ACCGCGCGCGCGCGCTGCTCTCGACAGTTCCCGAGCGCGCGCGCGCTCCATCG 702
Db 541 ACCGCGCGCGCGCGCTGCTCTCGACAGTTCCCGAGCGCGCGCGCGCTCCATCG 600
```

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD  
clade; Panicoideae; Andropogoneae; Zea.

1 (bases 1 to 893)

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,

Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,

Citek, R.W., Nunberg, A., Robbins, D. and Lakey, N.

Consortium for Maize Genomics

Unpublished (2002)

Other GSSs: OG3CN22TH

Contact: Cathy Whitelaw

7912 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..893

/organism="Zea mays"

/mol\_type="genomic DNA"

/strains="B73"

/db\_xref="taxon:4577"

/clone="ZMMBma0776C20"

/clone\_lib="ZM\_0.7\_1.5\_KB"

/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb

methylation filtered genomic DNA library"



RESULT 5		CG376770	730 bp	DNA	linear	GSS 26-AUG-2003
CG376770/c		OG3CN22TH ZM.0.7.1.5 KB	Zea mays	genomic clone	ZMMBma0776C20,	
LOCUS		genomic survey sequence.				
DEFINITION						
ACCESSION		CG376770	1 GI:34294037			
VERSION		CG376770	GSS.			
KEYWORDS		Zea mays				
SOURCE		Zea mays				
ORGANISM		Zea mays				
REFERENCE		Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nunberg,A., Robbins,D. and Lakey,N. (2002) Consortium for Maize Genomics Unpublished (2002) Other GSSs: OG3CN22TV				
AUTHORS		Contact: Cathy Whitelaw				
TITLE		TIGR				
JOURNAL		9712 Medical Center Drive, Rockville, MD 20850, USA				
COMMENT		Tel: 301-838-5843 Fax: 301-838-0208 Email: whitelaw@tigr.org Seq primer: TR Class: Sheared ends.				
FEATURES		Location/Qualifiers				
source		1. .730 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBma0776C20" /clone_lib="ZM.0.7.1.5 KB" /note="Vector: pBCKS-; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"				
ORIGIN		Query Match 25.7%; Score 711; DB 29; Length 730; Best Local Similarity 100.0%; Pred. No. 7.3e-63; Matches 711; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	530	CATCCCGCTGCTGCGAGCATGTATCGGCGCTGCTCGCGCGGACGCTCATGTGTCAGGT	589			
DB	730	CATCCCGCTGCTGCGAGCATGTATCGGCGCTGCTCGCGCGGACGCTCATGTGTCAGGT	671			
QY	590	CGTCTGCTCCAGTGCATCATCTGTACACGCTCATGCTCTTCTCTTCAGTACCGCGC	649			
DB	670	CGTCTGCTCCAGTGCATCATCTGTATACGCTCATGCTCTTCTCTTCAGTACCGCGC	611			
QY	650	CGCGCGCGCTGCTGCTTCGACAGTTTCCCGACGGCGCGCGCGCGCTCATGCTCTCTT	709			
DB	610	CGCGCGCGCTGCTGCTTCGACAGTTTCCCGACGGCGCGCGCGCTCATGCTCTCTT	551			
QY	710	CGCGCTTCAGTCCGACGCTGCTGCTCGCGCGGAGGAGTGTGAGCTCGAGGCGGAGCC	769			
DB	550	CGCGCTTCAGTCCGACGCTGCTGCTCGCGCGGAGGAGTGTGAGCTCGAGGCGGAGCC	491			
QY	770	CGACGCGCTGCGCGCGCGCGCGCTGCTCTCTCCCGCGCGGAGACCGCGCGCGGTCG	829			
DB	490	CGACGCGCTGCGCGCGCGCGCGCTGCTCTCTCCCGCGCGGAGACCGCGCGCGGTCG	431			
QY	830	CGTCAACGCTGCGAAGTCCACAGCTCGCGCTCGCGCGCGCTCGCGACTCGCATC	889			
DB	430	CGTCAACGCTGCGAAGTCCACAGCTCGCGCTCGCGCGCGCTCGCGACTCGCATC	371			
QY	890	CCAGACCATGACGCGCGCGCTGCTCCAACTCTCCGGGTGAGATCTACTCGTGCAGTC	949			
DB	370	CCAGACCATGACGCGCGCGCTGCTCCAACTCTCCGGGTGAGATCTACTCGTGCAGTC	311			
QY	950	GTCGCGCAACCCACCCCGCGCGCTTCAACCAACCGCGCTTCTTTCAACATCGT	1009			



LOCUS	CG722061	835 bp	DNA	linear	GSS 19-JUN-2003
DEFINITION	CGUIZ09TV ZM 0.7-1.5 KB Zea mays genomic clone ZMMBMa0452A18, genomic survey sequence.				
ACCESSION	CG722061				
VERSION	CG722061.1	GI:32126837			
KEYWORDS	GSS.				
SOURCE	Zea mays				
ORGANISM	Zea mays				
REFERENCE					
AUTHORS	Whitlaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Rietck,A., Frazer,C.M., Budiman,M.A., Bedell,J.A., Rohlfing,T., Citek,R.W., Nurnberg,A., Robbins,D. and Lakey,N.				
TITLE	Consortium for Maize Genomics				
JOURNAL	Unpublished (2002)				
COMMENT	Other GSSs: CGUIZ09TH Contact: Cathy Whitlaw TIGR				
	9712 Medical Center Drive, Rockville, MD 20850, USA Tel: 301-838-5843 Fax: 301-838-0208 Email: whitlaw@tigr.org Seq primer: TF Class: sheared ends.				
FEATURES					
source	1..835 /organism="Zea mays" /mol_type="genomic DNA" /strain="B73" /db_xref="taxon:4577" /clone="ZMMBMa0452A18" /clone_lib="ZM 0.7-1.5 KB" /note="Vector: pCSK+; Site 1: HincII; 0.7-1.5 kb methylation filtered genomic DNA library"				
ORIGIN					
Query Match	23.1%;	Score 639;	DB 29;	Length 835;	
Best Local Similarity	100.0%;	Pred. No. 1.3e-55;			
Matches	639;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	602	GTGCATCATCTGTGTACAGCTCATCTCTTCCTTCAGTACCGCGCGCGCGCGCT	661		
Db	1	GTGCATCATCTGTGTACAGCTCATCTCTTCCTTCAGTACCGCGCGCGCGCGCT	60		
Qy	662	CGTCCTCGACCACTTCCCGACGCGCGCGCGCTCCATCGTCTCTTCGCGTCGACTC	721		
Db	61	CGTCTCTGACCACTTCCCGACGCGCGCGCGCTCCATCGTCTCTTCGCGTCGACTC	120		
Qy	722	CGACGTGTCTTCGTTCGCGCAGGGGGGACGTTCGAGCTCGAGCCCGAGCCCGCGCTCGC	781		
Db	121	CGACGTGTCTTCGTTCGCGCAGGGGGGACGTTCGAGCTCGAGCCCGAGCCCGCGCTCGC	180		
Qy	782	CGGCGCGCGCGCGCTCTCTCCCGCGGGGACGCCGGGGGTGCGCTCACCGTGGC	841		
Db	181	CGGCGCGCGCGCGCTCTCTCTCCCGCGGGGACGCCGGGGGTGCGCTCACCGTGGC	240		
Qy	842	CAAGTCCACCAAGCTCGCGCTCCGAGGGCGGTGTCTCGCACTTCGCACTCCAGACCAATGCA	901		
Db	241	CAAGTCCACCAAGCTCGCGCTCCGAGGGCGGTGTCTCGCACTTCGCACTCCAGACCAATGCA	300		
Qy	902	GCCCGGTGTCTCAACCTCTCCGGGTGGAGATCTACTCGCTGAGTCTGTCCGCAACCC	961		
Db	301	GCCCGGTGTCTCAACCTCTCCGGGTGGAGATCTACTCGCTGAGTCTGTCCGCAACCC	360		
Qy	962	CACCCCGCGGGGTTCAGCTTTCACACGCGCACTTCTTCAACATCGTTCGCGCGCGCGC	1021		
Db	361	CACCCCGCGGGGTTCAGCTTTCACACGCGCACTTCTTCAACATCGTTCGCGCGCGCGC	420		
Qy	1022	CAAGGAGCGGAGAGAGCGGCGGGGACAGAGAGAGAGGCGCATGCGCGCGCGCGCGG	1081		
Db	421	CAAGGAGCGGAGAGAGCGGCGGGGACAGAGAGAGGCGCATGCGCGCGCGCGCGG	480		

125	CGAGTACGGCGCGACGACATACAGACAGGACGAGGAGCGCGCGCGGACAA	184
1316	GGCGGGCCGACGCTGTGCGAAGCTGGGGTCCAACTCGACCGCGAGCTGTACCCCAAGGA	1375
185	GGCGGGCCGACGCTGTGCGAAGCTGGGGTCCAACTCGACCGCGAGCTGTACCCCAAGGA	244
1376	CGACGGCGAGGGAGGGCGCGCGGGTGGCGATCCCGCGGAGCGTGTGATGACGGCGCT	1435
245	CGACGGCGAGGGGAGGACGGCGGC--GGCGATCCCGCGGAGCGTGTGATGACGGCGCT	304
1436	CATCCTCATCATGTGTGGAGGAAGCTGATCCGGAACCCCAACCTACTCCAGCCTCAT	1495
302	CATCCTCATCATGTGTGGAGGAAGCTGATCCGGAACCCCAACCTACTCCAGCCTCAT	361
1496	CGGGGTGCTGTGGTCCCTGGTCTCCTACAGTGGGGCATCAGATGCGAGGCATCATCGC	1555
362	CGGGGTGCTGTGGTCCCTGGTCTCCTACAGTGGGGCATCAGATGCGAGGCATCATCGC	421
1556	CCGGTCGATTTGATCTCTGTCGAGCGGGGTCTCCGGATGGCCATGTTTCAGCCTAGGCGCT	1615
422	CCGGTCGATTTGATCTCTGTCGAGCGGGGTCTCCGGATGGCCATGTTTCAGCCTAGGCGCT	481
1616	GTTTCATGGCGGTGACGCCGAGGATCATCGGTGCGGGAAACAGCTGGCGGCCATCGCGAT	1675
482	GTTTCATGGCGGTGACGCCGAGGATCATCGGTGCGGGAAACAGCTGGCGGCCATCGCGAT	541
1676	GGGGGTCCGGTTCGTGCGAGGCCCGCGGTTCATGGCGCGCGCCTCCATCGCGGTGGTCT	1735
542	GGGGGTCCGGTTCGTGCGAGGCCCGCGGTTCATGGCGCGCGCCTCCATCGCGGTGGTCT	601
1736	GGCGGGGTCTCTCTCCACATCGCATGTCAGGCTGCTCTGGCTCAGGGGATCGT	1792
602	GGCGGGGTCTCTCTCCACATCGCATGTCAGGCTGCTCTGGCTCAGGGGATCGT	658

RESULT	9
CC371169/c	
LOCUS	CC371169
DEFINITION	PURSK16TD ZM 0.6-1.0 KB Zea mays genomic clone ZMMBta52D07, genomic survey sequence.
	958 bp DNA linear GSS:19-MAY-2003

VERSION	CC371169.1	GI:30844786
KEYWORDS	GSS.	
SOURCE	Zea mays	
ORGANISM	Zea mays	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.	

REFERENCE  
1 (bases 1 to 958)  
Whitelaw C.A., Quackenbush, J., Van Aken, S., Utterback, T.,  
Resnick, A., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and  
Bennetzen, J.  
Maize Genomics Consortium  
Unpublished (2003)  
Other GSSs: FUSK16TB  
Contact: Cathy Whitelaw

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843  
Fax: 301-838-0208  
Email: [white@tigr.org](mailto:white@tigr.org)  
Seq primer: T<sub>9</sub>  
Class: sheared ends.

```

FEATURES
    source
        Location/Qualifiers
            1..958
                /organism="zebra mays"
                /mol_type="genomic DNA"
                /strain="B73"
                /db_xref="taxon:4577"
                /clone="ZMMETA523D07"
                /clone_lib="ZM 0.6-1.0 kb"
                /note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
                Cor selected genomic DNA library"

```

Query Match. 21.6%; Score 599; DB 28; Length 958;  
 Best Local Similarity 88.8%; Pred. No. 1.3e-51;  
 Matches 706; Conservative 0; Mismatches 5; Indels 84; Gaps 2;

QY	1608	CTAGGCTGTTCATGGCGCTCGACGCGAGGATCATCGCTGCGGGAAACAAGCTGGCGGCC	1667
DB	882	CGAGGCTGTTCATGGCGCTCGACGCGAGGATCATCGCTGCGGGAAACAAGCTGGCGGCC	823
QY	1668	ATCGCGATGGGCGTCCGGTTCGTGCGAGGCCCGCGGTCATGCGCGCGCTCCATCGCC	1727
DB	822	ATCGCGATGGGCGTCCGGTTCGTGCGAGGCCCGCGGTCATGCGCGCGCTCCATCGCC	763
QY	1728	GTGGTCTGCGCGCGCTCCCTCCACATCGCCATCG-	1764
DB	762	GTGGTCTGCGCGCGCTCCCTCCACATCGCCATCGTCCAGGTACGCGCCCTGTGGCTT	703
QY	1765	-	1764
DB	702	GAGGTAAACGGTGGAGTGTGTGATGGAAACAGGTAGGACCAACATTCGTTCCCTTCTTCT	643
QY	1765	TCCAGGCTGTCTGCTCCTCAGGGGATCGTCCGTTTCGTGTTTCGCAAGAGTACGGGTTTC	1824
DB	642	TCCAGGCTGTCTGCTCCTCAGGGGATCGTCCGTTTCGTGTTTCGCAAGAGTACGGGTTTC	583
QY	1825	ATCCGACATCTCGACACAGCGTATGTTCCCAATAACATCGCATGTTTTCATCACTTGCC	1884
DB	582	ATCCGACATCTCGACACAGCGTATGTTCCCAATAACATCGCATGTTTTCATCACTTGCC	523
QY	1885	ATAGTTAAACGGGAAAAAAGCAGAAAGCAATCGATGACGACGCACTGAATTCATATGAT	1944
DB	522	ATAGTTAAACGGGAAAAAAGCAGAAAGCAATCGATGACGACGCACTGAATTCATATGAT	463
QY	1945	TCATTACTATGATGGTGTCTTCATGCGAGTGCAGTCAAAGAACACCTAATAAGCACTGAT	2004
DB	462	TCATTACTATGATGGTGTCTTCATGCGAGTGCAGTCAAAGAACACCTAATAAGCACTGAT	403
QY	2005	CTAGGACAGCATCAGCATGATGATGCTTTCTTTCTCTGCAAACTCGCATTTCTTACT	2064
DB	402	CTAGGACAGCATCAGCATGATGATGCTTTCTTTCTCTGCAAACTCGCATTTCTTACT	343
QY	2065	ACAGAGTGCCTTCACATCATCCATCCAGATGATCATACAACACTACTGATGCATC-TTT	2123
DB	342	ACAGAGTGCCTTCACATCATCCATCCAGATGATCATACAACACTACTGATGCATCTTT	283
QY	2124	TTTTTTGATTCTGCTGACGCGTATCTTCGGGATGCTCATCGCTTCGCCCATCACTCTGG	2183
DB	282	TTTTTTGATTCTGCTGACGCGTATCTTCGGGATGCTCATCGCTTCGCCCATCACTCTGG	223
QY	2184	TCCTACTACTCTGCTCGGGCTGTGAGCTCTCTCGCTCGCTTCTTCAGCGTGGGGAAGG	2243
DB	222	TCCTACTACTCTGCTCGGGCTGTGAGCTCTCTCGCTCGCTTCTTCAGCGTGGGGAAGG	163
QY	2244	CGCATGCTGTGGTGTATCGGCCCCACATGAAATTTCTGATGAGCATTAGGATTACTAT	2303
DB	162	CGCATGCTGTGGTGTATCGGCCCCACATGAAATTTCTGATGAGCATTAGGATTACTAT	103
QY	2304	TAGCTTAGCGAAGAA 2318	
DB	102	TAGCTTAGCGAAGAA 88	

RESULT 10  
 BU080190  
 LOCUS  
 DEFINITION BU080190 610 bp mRNA linear EST 27-AUG-2002  
 946150G03.y1.946 - tassal primordium prepared by Schmidt lab Zea  
 mays cDNA, mRNA sequence.  
 ACCESSION BU080190  
 VERSION BU080190.1 GI:22521379  
 KEYWORDS EST.  
 SOURCE Zea mays  
 ORGANISM Zea mays

RESULT 10	BU080190	610 bp	mrna	linear	EST 27-AUG-2002
LOCUS	BU080190				
DEFINITION	946150G03.y1 946 - tassael primordium prepared by schmidt lab Zea mays CDNA, mRNA sequence.				
ACCESSION	BU080190				
VERSION	BU080190.1	GI:22521379			
KEYWORDS	EST.				
SOURCE	Zea mays				
ORGANISM	Zea mays				

Db	478	CCCGGTGCAATTCGATCCCTGTCGGACGCGGGTCTCGGGATGGCCATGTTTCAGCCTAGGCC	537
Qy	1615	TGTTCAATGGCGCTGCAGCGAGGATCATCGCTGCGGGAACAAGACTTGGCGGCATCGCGA	1674
Db	538	TGTTTCATGGCGCTGCAGCGGATCATCGCTGCGGGAACAAGACTTGGCGGCATCGCGA	597
Qy	1675	TGGG	1678
Db	598	TGGG	601
RESULT 11			
LOCUS	CA830783		
DEFINITION	1117011D11.y1.1117 - Unigene V from Maize mRNA linear EST 12-DEC-2002		
ACCESSION	CA830783		
VERSION	CA830783.1		
KEYWORDS	GI:26558548		
SOURCE	Zea mays		
ORGANISM	Zea mays		
REFERENCE	Eukarya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.		
AUTHORS	Walbot,V.		
TITLE	1 (bases 1 to 630)		
JOURNAL	Maine ESTs from various cDNA libraries sequenced at Stanford		
COMMENT	Unpublished (1999) Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 1117011 row: D column: 11.		
FEATURES	Location/Qualifiers		
source	1..630		
	/organism="Zea mays"		
	/mol_type="mRNA"		
	/db_xref="dbEST:946136B05.y1"		
	/db_xref="taxon:4577"		
	/clone_lib="1117 - Unigene V from Maize Genome Project"		
	/notes="This library represents the unique genes found in the fifth round of EST sequencing at Stanford University for the maize genome project. Sequences are present from library 946. Contigs were assembled using ZmDBAssembler and 2 representatives from each contig were selected for the Unigene set. All singlets were also selected."		
ORIGIN			
Query Match	20.7%;	Score 574.4;	DB 14; Length 630;
Best Local Similarity	96.7%;	Pred. No. 4.8e-49;	
Matches 609;	Conservative	0; Mismatches 16;	Indels 5; Gaps 2;
Qy	1136	GCAATGCTCGTCTGGAGTCTCCAGCGCTCGCCGCTGTCGAGCGCGCGCTGCACGT	1195
Db	6	GCAATGCTCGTCTGGAGTCTCCCTTTTCTCGCCCGTCTCCAGGC--GCGCGTGCACGT	63
Qy	1196	CTTCGGCGCGCGGGCGTCAACATGCGACGTCTTCGCAAGAGGCCAGGCGCTACGA	1255
Db	64	CTTCGGCGCGCGCGCTGACCATGCGACGTCTTCGCAAGAGGCCAGGCGCTACGA	123
Qy	1256	CGAGTACGGCGCGACGACTACAGCAGCAGGACGAGAGCGGCGCGCGGACAA	1315
Db	124	CGAGTACGGCGCGACGACTACAGCAGCAGGACGAGAGCGGCGCGGCGGACAA	183
Qy	1316	GGGCGGCGCGACGCTGTGCAAGCTGGGGTCCAACTCGACGCGCAGCTGATCCCCAAGGA	1375
Db	184	GGGCGGCGCGACGCTGTGCAAGCTGGGGTCCAACTCGACGCGCAGCTGATCCCCAAGGA	243
Qy	1376	CGACGCGCAGGGGAGGCGCGCGGTGGCGATGCGCGCGAGCGTATGACGGCT	1435

[illegible]

RESULT 12  
BU037419  
LOCUS  
DEFINITION  
627 bp mRNA linear EST 23-AUG-2002  
946138F12.Y1 946 - tassal primordium prepared by schmidt lab Zea  
mays cDNA, mRNA sequence.

ACCESSION	BU037419	
VERSION	BU037419.1	GI:22472939
KEYWORDS	EST.	
SOURCE	Zea mays	
ORGANISM	Zea mays	
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 637)	

AUTHORS	Walbot, V.
TITLE	Maize ESTs from various cDNA libraries sequenced at Stanford
JOURNAL	Unpublished (1999)
COMMENT	Contact: Walbot V Department of Biological Sciences Stanford University 855 California Ave, Palo Alto, CA 94304, USA Tel: 650 723 2227 Fax: 650 725 8221 Email: walbot@stanford.edu Plate: 946138 row: F column: 12.

```

FEATURES
source
Location/Qualifiers
1. 627
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="OH43"
/db_xref="taxon:4577"
/tissue_type="tassels"
/dev_stage="just after the transition from vegetative to
inflorescence development"
/lab_host="XLOLR"
/clone_lib="946 - tassel primordium prepared by Schmidt
lab"
/note="Organ: tassels; Vector: HybriZAP; Site 1: ECoRI;
Site 2: XhoI; George Chuck dissected immature tassels
between 1mm and 3mm. Sharon Stanfield prepared the cDNA
library in HybriZAP. Sample insert size range was 350 bp
to 3 Kb with a 1 Kb average."

```

## ORIGIN

	Query Match	20.5%;	Score 569,	DB 13;	Length 627;
	Best Local Similarity	97.0%;	Pred. No. 1,7e-48;		
	Matches 604;	Conservative	0;	Mismatches 10;	Indels 9; Gaps 2;
Qy	989	CGCGCACTTCTTCAACATCGTCGGCGCCGCCGCCAAGAGGAGCGGAGGAGCGCGGGGA	1048		
Db	14	CGCGCACTTCTTCAACATCGTCCTTTTCCGCCGCCAAGGAGGCGGAGGAGCGCGGGGA	73		
Qy	1049	CGAGGAGAGGGCGCATCGCGCGCGCGCGGAGGACACTCGCCGACGCGCAGACGCCGT	1108		
Db	74	CGAGGAGAGGGCGCAT-----GCGCGCGGAGGAGCACTCGCCGACGCGCAGGCCGT	127		
Qy	1109	CGCCGTGCGCGCAAGAGGAAGCACTGCACATGCTCTGTCTGGAGCTCCAGCGCTCGCC	1168		
Db	128	CGCCGTGCGCGCAAGAGGAAGCACTGCACATGCTCTGTCTGGAGCTCCAGCGCTCGCC	187		
Qy	1169	CGTGTCGAGCGCGCGCGCTGCACTCTTTCGCGCGCGCGCGCTGACCATGCCGACGT	1228		
Db	188	CGTCTCCGAGCGCGCGCGCTGCACTCTTTCGCGCGCGCGCGCTGACCATGCCGACGT	247		
Qy	1229	CCTCGCCAAAGGAGCCAGGCGCTTACGACAGATACGGCGCGGACGACTACAGCAGCAGGAC	1288		
Db	248	CCTCGCCAAAGGAGCCAGGCGCTTACGACAGATACGGCGCGGACGACTACAGCAGCAGGAC	307		
Qy	1289	GAAGAACGGGAGCGCGCGCGGAGCAAGGGCGGGCGACGCTGTGAGAGTGCGGGTCCAA	1348		
Db	308	GAAGAACGGGAGCGCGCGCGGAGCAAGGGCGGGCGACGCTGTGAGAGTGCGGGTCCAA	367		
Qy	1349	CTCGACGGCGCAGCTGTACCCCAAGACACGCGGAGGAGGGCGGCGCGGTGGCGAT	1408		
Db	368	CTCGACGGCGCAGCTGTACCCCAAGACACGCGGAGGAGGGCGGCGCGGC---GGCGAT	424		
Qy	1409	GCGCGCGGAGCGGTGATGACGGGCTCATCTCATCTGTTGTGAGAGAGCTGATCCG	1468		
Db	425	GCGCGCGGAGCGGTGATGACGGGCTCATCTCATCTGTTGTGAGAGAGCTGATCCG	484		
Qy	1469	GAACCCCAACACCTACTCCAGCGCTCATCGGGCTGCTGGTCCCTGGTCTCTCTACAGGTG	1528		
Db	485	GAACCCCAACACCTACTCCAGCGCTCATCGGGCTGCTGGTCCCTGGTCTCTCTACAGGTG	544		
Qy	1529	GCGCATCGAGTCCAGCGATCATCGCCGGTCTGATTTTGATCTCTGCGACGCGGTCT	1588		
Db	545	GCGCATCGAGTCCAGCGATCATCGCCGGTCTGATTTTGATCTCTGCGACGCGGTCT	604		
Qy	1589	CGGGATGGCCATGTTGAGCCTAG	1611		
Db	605	CGGGATGGCCATGTTGAGCCTAG	627		

RESULT 13					
CG456385					
LOCUS	794 bp	DNA	linear	GSS 17-SEP-2003	
DEFINITION	PUIKJ57TB8 ZM 0.6_1.0 KB Zea mays genomic clone ZMMeta0603118, genomic survey sequence.				

ACCESSION	
CG456385	
VERSION	
CG456385.1	GI:34841385
KEYWORDS	GSS.
SOURCE	Zea mays
ORGANISM	Zea mays Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 794)  
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,  
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and  
Bennetzen,J.  
Maize Genomics Consortium  
Unpublished (2003)  
Other\_GSSs: FUIKJ57DB  
Contact: Cathy Whitelaw  
TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 301-838-5843

Fax: 301-838-0208  
Email: whitelaw@igr.org  
Seq primer: TR  
Class: sheared ends.

## FEATURES

Location/Qualifiers  
source  
1..794  
/organism="Zea mays"  
/mol\_type="genomic DNA"  
/strain="B73"  
/db\_xref="taxon:4577"  
/clone="ZMMB7a0603118"  
/clone\_lib="ZM 0.6-1.0 kb"  
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high  
Cor selected genomic DNA library"

## ORIGIN

Query Match 20.5%; Score 568.4; DB 29; Length 794;  
Best Local Similarity 88.8%; Pred. No. 1.8e-48;  
Matches 673; Conservative 0; Mismatches 1; Indels 84; Gaps 2;  
QY 1608 CTAGGCTGTTTCATGGCGCTCGACCGAGAGATCATCGGTGCGGAAACAAGCTGGCGGCC 1667  
Db 37 CCAGGCTGTTTCATGGCGCTCGACCGAGAGATCATCGGTGCGGAAACAAGCTGGCGGCC 96  
QY 1668 ATCCGATGGCGCTCGGTTCGTCGAGGCCCGCGGTTCATGGCGCGCTTCATCGCC 1727  
Db 97 ATCCGATGGCGCTCGGTTCGTCGAGGCCCGCGGTTCATGGCGCGCTTCATCGCC 156  
QY 1728 GTCGGTCTGCGCGCTCTCTCCACATCGCCATCG----- 1764  
Db 157 GTCGGTCTGCGCGCTCTCTCCACATCGCCATCGCCATCGCGCCCTGTGGCTT 216  
QY 1765 ----- 1764  
Db 217 GACGTAAACGGTGGATGTGATGGAAACAGGTAGGACCAACATTCGTTTCCCTTCTTCT 276  
QY 1765 TCCAGGCTGCTCTGCTCAGGGGATCGTGGTTCGTTTCGCAAGAGAGTACGGCGTTC 1824  
Db 277 TCCAGGCTGCTCTGCTCAGGGGATCGTGGTTCGTTTCGCAAGAGAGTACGGCGTTC 336  
QY 1825 ATCCGATCGTGGACACAGCGTATGGTTCGAATAAATGTCATGTTTTCATCATCTGCC 1884  
Db 337 ATCCGATCGTGGACACAGCGTATGGTTCGAATAAATGTCATGTTTTCATCATCTGCC 396  
QY 1885 ATAGTTAACGGGAAAGAAAGCAGAAACGAATCGATGACGACCTGAATTCATATGAT 1944  
Db 397 ATAGTTAACGGGAAAGAAAGCAGAAACGAATCGATGACGACCTGAATTCATATGAT 456  
QY 1945 TCATTACTAATGATGGTGTGTTTCATGACGTGACGTCAAGAACCACTAATAAGCACTGAT 2004  
Db 457 TCATTACTAATGATGGTGTGTTTCATGACGTGACGTCAAGAACCACTAATAAGCACTGAT 516  
QY 2005 CTAGACAGCATCAGCATGATGATGCTGTTTCTCCTGACAACTCTGCAATCTGATTTTACT 2064  
Db 517 CTAGACAGCATCAGCATGATGATGCTGTTTCTCCTGACAACTCTGCAATCTGATTTTACT 576  
QY 2065 ACACAGTGTGCTTCACTCATCCATCCAGATGATCATAACACTACTGATGATC-TTT 2123  
Db 577 ACACAGTGTGCTTCACTCATCCATCCAGATGATCATAACACTACTGATGATC-TTTT 636  
QY 2124 TTTTGTGATTCGTGACGCGTATCTTCGGATGCTCATGCTCTGCCCATCACCCTGG 2183  
Db 637 TTTTGTGATTCGTGACGCGTATCTTCGGATGCTCATGCTCTGCCCATCACCCTGG 696  
QY 2184 TCTACTACATCTTGTGGGGGTGTGAGCCTCTCTCGCTCGCTTCTTCAGCGGTGCGGAAGG 2243  
Db 697 TCTACTACATCTTGTGGGGGTGTGAGCCTCTCTCGCTCGCTTCTTCAGCGGTGCGGAAGG 756  
QY 2244 GGCATGTGTGGTGTATCGGCCCAATGAATTTCT 2281  
Db 757 GGCATGTGTGGTGTATCGGCCCAATGAATTTCT 794

## RESULT 14

CG317974

## LOCUS

CG317974 776 bp DNA linear GSS 26-AUG-2003  
OGWGA91TV ZM\_0.7-1.5 KB Zea mays genomic clone ZMMBma0570013,  
genomic survey sequence.

## DEFINITION

CG317974

## ACCESSION

CG317974

## VERSION

CG317974.1

## KEYWORDS

GSS.

## SOURCE

Zea mays

## ORGANISM

Zea mays

## REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Unpublished (2002)

Other\_GSSs: OGWGA91TV

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@igr.org

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..776

/organism="Zea mays"

/mol\_type="genomic DNA"

/strain="B73"

/db\_xref="taxon:4577"

/clone="ZMMBma0570013"

/clone\_lib="ZM 0.7-1.5 KB"

/note="Vector: pBCSK-; Site 1: HincII; 0.7-1.5 kb  
methylation filtered genomic DNA library"

## ORIGIN

Query Match 20.3%; Score 561; DB 29; Length 776;  
Best Local Similarity 87.4%; Pred. No. 1e-47;  
Matches 679; Conservative 0; Mismatches 0; Indels 98; Gaps 2;  
QY 795 GTCTCTCTCCCGCGGGGACCGCCGGGGGGTGCCTGTCACCGTCCGCAAGTCCACGAC 854  
Db 1 GTCTCTCTCCCGCGGGGACCGCCGGGGGGTGCCTGTCACCGTCCGCAAGTCCACGAC 60  
QY 855 TCGCGTCCGAGCGCGTGTCTGCACTCGCACTCCGAGACCATCGAGCCCGCTGTCTCC 914  
Db 61 TCGCGTCCGAGCGCGTGTCTGCACTCGCACTCCGAGACCATCGAGCCCGCTGTCTCC 120  
QY 915 AACCTCTCCGGCGTGGAGATCTACTCGCTGCACTGCTCGCAACCCACCCCGCGCGG 974  
Db 121 AACCTCTCCGGCGTGGAGATCTACTCGCTGCACTGCTCGCAACCCACCCCGCGCGG 180  
QY 975 TCCAGTTCAACCAACCGGACCTTCTTCAACATCTGTCGGCGCGCCGCCCAAGGAGCGGA 1034  
Db 181 TCCA-CTTCAACCAACCGGACCTTCTTCAACATCTGTCGGCGCGCCGCCCAAGGAGCGGA 239  
QY 1035 GGAGCGCGGGGAGCGAGGAGAGGCGCATGCGCGCGCGGGCGGAGGACACTCGCCG 1094  
Db 240 GGAGCGCGGGGAGCGAGGAGAGGCGCATGCGCGCGCGGGCGGAGGAGGACACTCGCG 299  
QY 1095 CAGCGCGAGCGCTCGCGTCCGCCCAAGAGAGGACCTTGCACATGCTCTGTGAGC 1154  
Db 300 CAGCGCGAGCGCTCGCGTCCGCCCAAGAGAGGACCTTGCACATGCTCTGTGAGC 359  
QY 1155 TCCAGCGCTCGCGCTGTTCGAGCGCGCGCGCTTTCGGCGCGCGCGCT 1214  
Db 360 TCCAGCGCTCGCGCTGTTCGAGCGCGCGCGCTTTCGGCGCGCGCGCT 419  
QY 1215 GACCATGCCGCGCTCTCGCCAA----- 1237

```

Db      420  GACCATGCCGAGCTCTCGCCAAAGGTTGGTAACAATGCGTGATCTTTTGGTTGATTTC 479
QY      1238  ----- 1237

Db      480  CTAACCAAAATGTTGTCTGCTGCAFTTTTATTTGTACCAATGGGTGTGCACAACAAGC 539
QY      1238  AGGAGCCAGGCTACGACGAGTACGGCGGACGACTACAGCAGCAGGAGCAAGAAGCG 1297
Db      540  AGGAGCCAGGCTACGACGAGTACGGCGGACGACTACAGCAGCAGGAGCAAGAAGCG 599
QY      1298  GAGCGCGCGCGGACAAAGGGCGGCGACGCTGTCGAAGCTGGGGTCCAACTCGACGGC 1357
Db      600  GAGCGCGCGCGGACAAAGGGCGGCGACGCTGTCGAAGCTGGGGTCCAACTCGACGGC 659
QY      1358  CGAGCTGTACCCCAAGGACGAGCGGAGGAGGCGCGCGGTGGCGATGCCCGCGGC 1417
Db      660  CGAGCTGTACCCCAAGGACGAGCGGAGGAGGCGCGCGGTGGCGATGCCCGCGGC 719
QY      1418  GAGCGTGATGACGCGGCTCATCTTCATCATGTGTGGAGGAAGCTGATCCGGAACCC 1474
Db      720  GAGCGTGATGACGCGGCTCATCTTCATCATGTGTGGAGGAAGCTGATCCGGAACCC 776

RESULT 35
CA831110 1117015F07.y1 1117 - Unigene V from Maize linear EST 12-DEC-2002
LOCUS    CDNA, mRNA sequence.
DEFINITION
ACCESSION CA831110
VERSION   CA831110.1 GI:26558875
KEYWORDS  EST.
SOURCE    Zea mays
           Eukarya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
           clade; Panicoideae; Andropogoneae; Zea.
           1 (bases 1 to 604)
           Walbot.V.
           Maize ESTs from various cDNA libraries sequenced at Stanford
           University
           Unpublished (1999)
           Contact: Walbot V
           Department of Biological Sciences
           Stanford University
           855 California Ave, Palo Alto, CA 94304, USA
           Tel: 650 723 2227
           Fax: 650 725 8221
           Email: walbot@stanford.edu
           Plate: 1117015 row: F column: 07.
FEATURES
     source
         1..604
             /organism="Zea mays"
             /mol_type="mRNA"
             /db_xref="dbEST:946139F12.y1"
             /db_xref="taxon:4577"
             /clone_lib="1117 - Unigene V from Maize Genome Project"
             /note="This library represents the unique genes found in
the fifth round of EST sequencing at Stanford University
for the maize genome project. Sequences are present from
library 946. Contigs were assembled using ZmBAssembler
and 2 representatives from each contig were selected for
the Unigene set. All singlets were also selected."
ORIGIN
Query Match 20.2%; Score 559.8; DB 14; Length 604;
Best Local Similarity 97.4%; Pred.No.1.5e-47;
Matches 593; Conservative 0; Mismatches 7; Indels 9; Gaps 2;

QY      989  CGCGGACTTCTTCAACATCGTCGGCGCCCGCCCAAGGAGCGGAGGAGCGCGGGGA 1048
Db      5    CGCGGACTTCTTCAACATCGTCGGTCGCCGCCCAAGGAGCGGAGGAGCGCGGGGA 64

```

```

QY      1049  CGAGGAGAAAGGCGCATCGCGGGCGGAGGACACTCGCCGACAGCCAGCCCGT 1108
Db      65    CGAGGAGAAAGGCGCAT-----GCGGCGGCGAGGACACTCGCCGACAGCCAGCCCGT 118

QY      1109  CGCGTGCCTCGGCAAGAGAAAGCACTGCACATGCTCTGAGCTCCAGCGCTCGCC 1168
Db      119  CGCGTGCCTCGGCAAGAGAAAGCACTGCACATGCTCTGAGCTCCAGCGCTCGCC 178

QY      1169  CGTGTCCGAGCGCGCGCGCTGCACGTCTTTCGCGCGCGCGCGCTGACATGCGACGT 1228
Db      179  CGTGTCCGAGCGCGCGCGCTGCACGTCTTTCGCGCGCGCGCGCTGACATGCGACGT 238

QY      1229  CTTTCGCAAAAGAGGCCAGGCCCTACGACGATACGGGCGCGAGCTACAGCAGCAGGAC 1288
Db      239  CTTTCGCAAAAGAGGCCAGGCCCTACGACGATACGGGCGCGAGCTACAGCAGCAGGAC 298

QY      1289  GAGAACGCGGAGCGCGCGCGCAAGAGGCGCGCGCGCTGTGAAAGCTGGGGTCCAA 1348
Db      299  GAGAACGCGGAGCGCGCGCGCGCAAGAGGCGCGCGCGCTGTGAAAGCTGGGGTCCAA 358

QY      1349  CTCGACGCGCGAGCTGTACCCCAAGGACGACGCGGAGGAGGCGCGCGCGTGGCGAT 1408
Db      359  CTCGACGCGCGAGCTGTACCCCAAGGACGACGCGGAGGAGGCGCGCGCGCGCGC---GCGCAT 415

QY      1409  GCGCGCGCGAGCGCTGATGACGCGGCTCATCTCATATGCTGTGAGGAAGCTGATCG 1468
Db      416  GCGCGCGCGAGCGCTGATGACGCGGCTCATCTCATATGCTGTGAGGAAGCTGATCG 475

QY      1469  GAAACCCCAACACTACTTCCAGGCTCATCGCGCTCGTCTGCTCCCTGCTCTCTACAGTG 1528
Db      476  GAAACCCCAACACTTACTTCCAGGCTCATCGCGCTCGTCTGCTCCCTGCTCTCTACAGTG 535

QY      1529  GGCATCGAGATGCCAGCGATCATCGCCCGGTCGATTTTCGATCCCTGTTCGAGCGCGGTCT 1588
Db      536  GGCATCGAGATGCCAGCGATCATTCGCCCGGTCGATTTTCGATCCCTGTTCGAGCGCGGTCT 595

QY      1589  CGGGATGGC 1597
Db      596  CGGGATGGC 604

```

Search completed: March 3, 2004, 06:17:59  
Job time : 6909 secs



QY 385 TGCCGCGGACACGCTGCGAGAGTGGCGCTCTCTGCGGTGCTGGCGGTGGCGCTTCCCGCG 444  
Db |||||  
729 TCCCGCTCTGCTGCTTCCCGCTCTCTCCCGTCTCTGCTCTCCCGTCTCTCCCGTCTCTCG 788  
QY 445 GCCTCTCTCCCGCGGCGCTGGCTGAGTGGAGATACGCTCTTCTCCCTCTCCA 504  
Db |||||  
789 TCTCTCCCGTCTCTCCCGTCTCTGCTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCC 848  
QY 505 CGTCTCCCAACAGCTGCTCATGGGATCCCGCTCTGCGAGGATGTAAGCGCGGTCTGT 564  
Db |||||  
849 CGGTCTCCCGCTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCT 908  
QY 565 CGGCGGCGAGCTCATGGTTCAGTCTGCTCTCCAGTGCATCATCTGGTACAGCTCA 624  
Db |||||  
909 TCTCTGCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCT 968  
QY 625 TGCT 680  
Db |||||  
969 TGTCT 1028  
QY 681 GACGCGCGCGCGCTGCT 740  
Db |||||  
1029 TCGGCGCGCGCGCT 1088

## RESULT 2

US-09-130-114-2  
; Sequence 2, Application US/09130114  
; Patent No. 5976807  
; GENERAL INFORMATION:  
; APPLICANT: Horlick, Robert A.  
; APPLICANT: Damaj, Bassem B.  
; APPLICANT: Robbins, Alan K.  
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes  
; TITLE OF INVENTION: From Multiple Transfected Episomes  
; FILE REFERENCE: 0867/1D903US1  
; CURRENT APPLICATION NUMBER: US/09/130,114  
; CURRENT FILING DATE: 1998-08-06  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 2  
; LENGTH: 1931  
; TYPE: DNA  
; ORGANISM: EBNA  
US-09-130-114-2

Query Match 4.1%; Score 114.8; DB 2; Length 1931;  
Best Local Similarity 48.5%; Pred. No. 1.3e-12;  
Matches 349; Conservative 0; Mismatches 367; Indels 4; Gaps 1;

QY 25 CCACT 84  
Db |||||  
369 CCGTCT 428  
QY 85 CTTCT 144  
Db |||||  
429 TCGTCT 488  
QY 145 CGTCCCGGCAAGCCATCCGCGCCATGATACACCGCGTGGACCTCTACCACTGCTGA 204  
Db |||||  
489 TCCTCCCGGCT 548  
QY 205 CGGCGGTGGTGGCTGTAGTGGCCATGACGTGGAGTACGCTCCGCTCCGCTGGTGGC 264  
Db |||||  
549 CGGTCT 608  
QY 265 GCATCTTTCAGCGCGGACAGTCTCGGGATCAACCGCTTCTGGCGCTCTTTCGCGGTGC 324  
Db |||||  
609 TCGTCT 668  
QY 325 CGTCT 384  
Db |||||  
669 CGGTCT 728

QY 385 TGCCGCGGACACGCTGCGAGAGTGGCGCTCTCTGCGGTGCTGGCGGTGGCGCTTCCCGCG 444  
Db |||||  
729 TCCCGCTCTGCTGCTTCCCGTCTCTCCCGTCTCTGCTCTCCCGTCTCTCCCGTCTCTCG 788  
QY 445 GCCT 504  
Db |||||  
789 TCTCTCCCGTCTCTCCCGTCTCTGCTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCC 848  
QY 505 CGTCTCCCAACAGCTGCTCATGGGATCCCGCTCTGCGAGGATGTAAGCGCGGTCTGT 564  
Db |||||  
849 CGGTCTCCCGCTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCT 908  
QY 565 CGGCGGCGAGCTCATGGTTCAGTCTGCTCTCCAGTGCATCATCTGGTACAGCTCA 624  
Db |||||  
909 TCTCTGCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCTCTCCCGTCT 968  
QY 625 TGCT 680  
Db |||||  
969 TGTCT 1028  
QY 681 GACGCGCGCGCGCTGCT 740  
Db |||||  
1029 TCGGCGCGCGCGCT 1088

## RESULT 3

US-09-103-840A-2  
; Sequence 2, Application US/09103840A  
; Patent No. 6294328  
; GENERAL INFORMATION:  
; APPLICANT: FLEISCHMAN, Robert D.  
; APPLICANT: WHITE, Owen R.  
; APPLICANT: FRASER, Claire M.  
; APPLICANT: VENTER, John C.  
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM  
; TITLE OF INVENTION: TUBERCULOSIS  
; FILE REFERENCE: 24366-20007.00  
; CURRENT APPLICATION NUMBER: US/09/103,840A  
; CURRENT FILING DATE: 1998-06-24  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 4403765  
; TYPE: DNA  
; ORGANISM: Mycobacterium tuberculosis  
; FEATURE:  
; OTHER INFORMATION: CDC 1551  
; OTHER INFORMATION: "n" bases at various positions throughout the sequence  
; OTHER INFORMATION: represent a, t, c or g  
US-09-103-840A-2

Query Match 3.6%; Score 100.2; DB 3; Length 4403765;  
Best Local Similarity 46.3%; Pred. No. 5.4e-09;  
Matches 368; Conservative 0; Mismatches 423; Indels 4; Gaps 1;

QY 643 ACCGCGCGCGCGCGCTCTCTCTGACCACTTCCCGGACGCGCGCGCGCGCTTCCATCG 702  
Db 335128 ATCCGCGCGGAGCGCTGCGCACTTCCCGCGCGCGCGCGCGCGCGCGCGCGCG 335187  
QY 703 TCTCTTCCCGGTGATCTCCGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 762  
Db 335188 ACCGTGCGCGGTGTCGCCCGCGCGCGCGCTTGCCTGCGCGCGCGCGCGCGCGCG 335247  
QY 763 CCGAGCGGAGCGGTGTCGCCCGCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCT 822  
Db 335248 GCGCGCGCGGAGCGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335307  
QY 823 GGGTGCAGTCACTCGTCCGCAAGTCCACCACTGCGCTCTGAGCGCGCGCGCGCGCG 882  
Db 335308 CCGCTATGCGCGCGCGCGGTGAGAGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 335367  
QY 883 CGCACTCCAGACCATGACGCGCGGTGCTCTCAACCTCTCCGCGGTGAGATCTACTCGC 942



```
Query Match 3.6%; Score 99.2; DB 3; Length 4411529;
Best Local Similarity 46.4%; Pred. No. 8.3e-09;
Matches 359; Conservative 0; Mismatches 413; Indels 2; Gaps 1;

QY 642 TACCGCGCGCGCGCGCTCTCTCGACCAAGTTCCTCCGACGGCGCGCGCTCCATC 701
DB 3930731 TACGGTGATCGCGGTCAACCTACGCGACGCGCGCCAGCGCGCGCGCTGCC 3930672
QY 702 GTCCTCTTCGCGGTGCACTCCGACGTCTCGCTCGCGACGGGGAGTGGAGTCGAG 761
DB 3930671 GCGGTGCGCGCGTTCGCGCGGTACCGCGGTGCGCGGTGCGCGGTGCGCGGTTC 3930612
QY 762 GCGGAGCCCGAGCGGTTCGCGCGCGCGCGCGCTCTCTCCGCGGGGGAGCGCGGG 821
DB 3930611 GCGGTCTCTGCTCGCGCGCGCGGTGCGCGCAATTGCGCGCGGTTCGCGCGAA 3930552
QY 822 CGGCTGCGCGTCACTCGGTGCGCAAGTTCACCACTCGCGTTCGAGGCGCGGTGTCGAC 881
DB 3930551 CCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930492
QY 882 TCGCACTCCAGACCATGACGCGCGCGGTGTCTCAACCTCTCCGCGGTGGAGATCTACTCG 941
DB 3930491 GCGCGCGCGGTAAACGCGGTTCGCGCGGTTCGCGCGGTTCGCGCGCGCGCGCG 3930432
QY 942 CTGCACTGCTGCGCAACCCGACCGCGCGCGGTTCAGTTCACCAACGCGCACTTCTTC 1001
DB 3930431 GCGCGGTGCGCGCAACCGCGCAATTGCTGCGGTGCGCGCGCGCGCGCGCGCGGT 3930372
QY 1002 AACATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1061
DB 3930371 GCGCGGTTCGCGCAATGCGCAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930312
QY 1062 GATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1121
DB 3930311 GTTTCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930252
QY 1122 AAGAGGAGGACTGACATGCTCTGCTGAGCTCCAGCGCGTCCGCGCGGTTCGCGCGCG 1181
DB 3930251 GCGGAACCGCGGTGATGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930192
QY 1182 GCGCGCGTGCAGTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1241
DB 3930191 GCGCGCGTGCAGTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930132
QY 1242 GCGCGCGTGCAGTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1301
DB 3930131 GCGGTGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930072
QY 1302 GCGCGCGCGCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1361
DB 3930071 GGTGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3930014
QY 1362 CTGTATCCCAAGACGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1415
DB 3930013 CGGTGCGCGCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3929960
```

```
RESULT 6
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 629432B
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: PRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2
```

```
Query Match 3.5%; Score 98; DB 3; Length 4403765;
Best Local Similarity 47.1%; Pred. No. 1.4e-08;
Matches 369; Conservative 0; Mismatches 410; Indels 5; Gaps 2;

QY 635 CTTCGAGTACCGCGCGCGCGCGCTCTCTCGACCAAGTTCCTCCGACGGCGCGCGCGCG 694
DB 3924520 CTACGGTGATCGCGGTCAACCTACGCGACGCGCGCGCGCGCGCGCGCGCGCGCG 3924461
QY 695 GTCATGCTCTCTTCGCGGTGCACTCCGACGTCGTCGTCGCGAGGGGACGTCGA 754
DB 3924460 CGCGCGGTGCGCGCGGTTCGCGCGGTACCGCGGTGCGCGGTGCGCGGTTCGCGGT 3924401
QY 755 GCTCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 814
DB 3924400 CGCGGTCTCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3924341
QY 815 CGCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 871
DB 3924340 ACCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3924281
QY 872 GTGCTGCACTCGCACTCCAGACCATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 931
DB 3924280 CGCGCGCGCGGTAAAGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCG 3924221
QY 932 GATCTACTCTGCTGCTGCTGCGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 991
DB 3924220 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3924161
QY 992 CGATCTTCTTCAACATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1051
DB 3924160 AGCGCGCGGTGCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3924101
QY 1052 GGAGAAAGGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1111
DB 3924100 CGGTGCGCGGTTCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3924041
QY 1112 CGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1171
DB 3924040 CGCGCGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3923981
QY 1172 GTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1231
DB 3923980 CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3923921
QY 1232 CGCGAAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1291
DB 3923920 CGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3923861
QY 1292 GAACGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1351
DB 3923860 CGTTGCGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCGGTGCGCG 3923803
QY 1352 GACGCGCGCGGTACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1411
DB 3923802 GCGCGCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3923743
QY 1412 GCGC 1415
DB 3923742 GCGC 3923739
```

RESULT 7





INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2580 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-359-081-2

Query Match 3.4%; Score 92.8; DB 4; Length 2580;  
Best Local Similarity 47.1%; Pred. No. 1.7e-08;  
Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;

122 CGCGACCCCTCCCTCCCTCCCGGCTCCCGGACAGCCATCCCGGCGCCATGATCAACGC 181  
1444 CGGCGGCTCCACTACCTCCCTCGACCGGCTCCACTACCTCCCTCGACCGGCGCTCCA 1385  
182 GCTGGACCTTACCACTGCTGACGGCGGTGTCGCTGACGTGACGATGAGCTGGC 241  
1384 CTGCTCTCTGACCGCGGCTCCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1325  
242 GTACGGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301  
1324 CTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265  
302 CTTCGTGCGGCTCTTCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 361  
1264 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1205  
362 CTTCGCAATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421  
1204 CTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1145  
422 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 481  
1144 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1085  
482 CATCAGCT 529  
1084 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1025  
530 CATCCGCTGCTGCGAGCATATGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589  
1024 CTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 965  
590 GCTGCTCTCAGTGATCATGTGATACAGCTCATGCTCTTCTCTCTCTCTCTCTCTCTCT 649  
964 CTGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 905  
650 CGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709  
904 CTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 847  
710 CGCGTGCAGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769  
846 CTGCGCT 787  
770 CGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829  
786 CTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 727  
830 GGTCACTGCGCAAGTCCACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 877  
726 CCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 679

RESULT 11  
US-09-130-114-1  
Sequence 1, Application US/09130114  
Patent No. 5976807  
GENERAL INFORMATION:  
APPLICANT: Horlick, Robert A.

APPLICANT: Damaj, Bassam B.  
APPLICANT: Robbins, Alan K.  
TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes  
TITLE OF INVENTION: From Multiple Transfected Episomes  
FILE REFERENCE: 0867/1D903US1  
CURRENT APPLICATION NUMBER: US/09/130,114  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 5452  
TYPE: DNA  
ORGANISM: VEBNA  
US-09-130-114-1

Query Match 3.4%; Score 92.8; DB 2; Length 5452;  
Best Local Similarity 47.1%; Pred. No. 2.1e-08;  
Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;

122 CGCGACCCCT 181  
1361 CGGCGGCTTCCACTACCTCTCTGACCGGCTCCACTACCTCTCTGACCGGCGCTCCA 1420  
182 GCTGGACCTTACCACTGCTGACGGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 241  
1421 CTGCTCTCTGACCGGCGCTCCACTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCTCT 1480  
242 GTACGGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 301  
1481 CTGCTCTCTGCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1540  
302 CTTCGTGGGCTCTTCGCGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361  
1541 CTCTCTGCGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1600  
362 CTTCGCAATGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 421  
1601 CTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1660  
422 GCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 481  
1661 CTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1720  
482 CATCAGCT 529  
1721 CTCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1780  
530 CATCCGCTGCTGCGAGCATATGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 589  
1781 CTGCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1840  
590 GCTGCT 649  
1841 CTGCGGCT 1900  
650 CGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709  
1901 CT 1958  
710 CGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769  
1959 CTGCGGCT 2018  
770 CGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 829  
2019 CT 2078  
830 GGTCACTGCGCAAGTCCACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 877  
2079 CCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2126

RESULT 12

US-09-647-344A-14  
; Sequence 14, Application US/09647344A  
; Patent No. 6586180  
; GENERAL INFORMATION:  
; APPLICANT: Ruffner, Duane E.  
; APPLICANT: Pierce, Michael L.  
; APPLICANT: Chen, Zhidong  
; TITLE OF INVENTION: Directed Antisense Libraries  
; FILE REFERENCE: T6678.PCT.US  
; CURRENT APPLICATION NUMBER: US/09/647,344A  
; CURRENT FILING DATE: 2000-12-04  
; PRIOR APPLICATION NUMBER: PCT/US99/06742  
; PRIOR FILING DATE: 1999-03-28  
; NUMBER OF SEQ ID NOS: 50  
; SEQ ID NO 14  
; LENGTH: 8705  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: pshuttle  
US-09-647-344A-14

Query Match 3.4%; Score 92.8; DB 4; Length 8705;  
Best Local Similarity 47.1%; Pred. No. 2.4e-08;  
Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;  
QY 122 CGCGGACCCCTCTCTCCCTCCCGCGTCCCGGACCAAGCCATCCGGGCGCATGATCACCGC 181  
Db 7226 CGGGGGCTCCACTACTCTCTGACCCCGGCTCCACTACTCTCTGACCCCGGCTCCA 7285  
QY 182 GCTGGACCTTACACGTGTGACGGGGTGTGGCGGTGTAGTGGCCATGACGCTGGC 241  
Db 7286 CTGCTCTCTGACCCCGGCTCCACCTCTCTGCTCTGCTCTCTGCTCTCTCTCTCTCT 7345  
QY 242 GTACGGCTCGCTCGCTGTGGGCGCATCTTACGCGGACCAAGTCTCCGGGATCAACGG 301  
Db 7346 CTGCTCTCTGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7405  
QY 302 CTTCGTGGCGCTCTTCGCGGTGCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361  
Db 7406 CTCTCTGCGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7465  
QY 362 CTTCGGCATGAACCTGGCTCTGCGGCGGCGGACGCTGACGAGGTGGCGTCTCGC 421  
Db 7466 CT 7525  
QY 422 GCTGCTGGCGCTGCGCTCTCCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 481  
Db 7526 CTTCGCGCTCTCTGCGCT 7585  
QY 482 CATCAGCT 529  
Db 7586 CT 7645  
QY 530 CATCCCGCTCTGCGAGGCTATGACGGCGGTGTGCGGCGGCGACGCTCATGGTCCAGGT 589  
Db 7646 CTTCGCGCTCTCTGCGCT 7705  
QY 590 GCTGCTCTCTGAGTGATCATGTGTACAGCTCATGCTCTCTCTCTCTCTCTCTCTCTCT 649  
Db 7706 CTTCGCGCTCTCTGCGCT 7765  
QY 650 CGCGCGGCGCT 709  
Db 7766 CT 7823  
QY 710 CCGGCTGACATCCGACGT 769  
Db 7824 CTGCGCT 7883  
QY 770 CGAGGGGTGTGCGGCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 829  
Db 7884 CT 7943

QY 830 CTTACCGTGGCGCAAGTCCACAGCTCGGCTCCGAGCGCGCTGCTC 877  
Db 7944 CCCT 7991  
RESULT 13  
US-08-910-647-1/c  
; Sequence 1, Application US/08910647  
; Patent No. 6251433  
; GENERAL INFORMATION:  
; APPLICANT: Zuckermann et al.  
; TITLE OF INVENTION: Compositions and Methods for  
; TITLE OF INVENTION: Polynucleotide Delivery  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Chiron Corporation  
; STREET: 4560 Horton Street  
; CITY: Emeryville  
; STATE: California  
; COUNTRY: U.S.A.  
; ZIP: 94608-2916  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/910,647  
; FILING DATE:  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fujita, Sharon M.  
; REGISTRATION NUMBER: 38,459  
; REFERENCE/DOCKET NUMBER: 1218.002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (510) 923-2706  
; TELEFAX: (510) 655-3542  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9600 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-08-910-647-1

Query Match 3.4%; Score 92.8; DB 3; Length 9600;  
Best Local Similarity 47.1%; Pred. No. 2.5e-08;  
Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;  
QY 122 CGCGGACCCCT 181  
Db 1490 CGGGGCTCTCACTACTCTCTCTGACCCCGGCTCCACTACTCTCTGACCCCGGCTCCA 1431  
QY 182 GCTGACCTCTTACACAGTGTGACGGGGTGTGGCTGTGCTGTGCTGTGCTGTGCTGTG 241  
Db 1430 CTGCTCTCTGACCCCGGCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1371  
QY 242 GTACGGCT 301  
Db 1370 CTGCTCTCTGCT 1311  
QY 302 CTTGTGGCGCTCTTGGCGGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 361  
Db 1310 CTCTCTGCGCT 1251  
QY 362 CTTGCGCATGAACCTGGCTCTCTGCGGCGGCGACAGCTGCAGAGGTGGCGCTCTCTCG 421  
Db 1250 CTTCT 1191  
QY 422 GCTGTGGCGCTGGCT 481



COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: 19920518  
FILING DATE: 19920518  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 755.1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 10596 bases  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-884-811-15

Query Match 3.4%; Score 92.8; DB 1; Length 10596;  
Best Local Similarity 47.1%; Pred. No. 2.5e-08;  
Matches 362; Conservative 0; Mismatches 392; Indels 14; Gaps 2;  
QY 122 CGCGGACCGCTCTCCCTCCGCGCTCCCGGCGTCCCGACAGCCATCGCGGGCATGATCACCGC 181  
DB 3024 CGCGGCGCTCCATCTACTCTGACCCCGGCTCCACTACCTCCTCGACCCCGGCTCCA 2965  
QY 182 GCTGAGACCTTACACGCTGTACGCGGTGTGCGCTGTACGTGGCGCATGACGCTGCG 241  
DB 2964 CTGCTCTCTGACCCCGGCTCCACCTCTGCTCTGCGCCCTCTGCTCTGCGCCCTCT 2905  
QY 242 GTACGGCTCGGTGCGGTGGGATCTTACGCGGACGAGTGTGCGGATCAACGG 301  
DB 2904 CTTGCTCTGCGCCCTCTGCGCCCTCTGCTGCTGCGCCCTCTGCGCCCTCTGCT 2845  
QY 302 CTTGCTGGGCTCTTCCGCGTGGCTCTCTCTCTTCCACTTCACTCCACCAAGAGCC 361  
DB 2844 CTTCTGCGCCCTCTGCTGCTGCGCCCTCTGCGCCCTCTCTGCTCTGCGCCCTCT 2785  
QY 362 CTTGCGCATGAACCTGGCTTCTGCGCCGCGACACGCTGACAGGTGGCGCTCTGCGC 421  
DB 2784 CTTCTCTGCTCTGCTGCGCCCTCTGCTGCTGCTGCGCCCTCTGCTGCGCCCTCT 2725  
QY 422 GCTGCTGGGCTGGCTCCGCGGCTCTCTCTCCGCGGCTGCGGCTGCGGCTGCTGAG 481  
DB 2724 CTTGCGCCCTCTGCTGCTGCGCCCTCTGCTGCTGCTGCGCCCTCTGCTGCTGCGC 2665  
QY 482 CATCAGCTCTTCTCCCTC-----TCCAGCTCTCCCAACACGCTGCTCATGGG 529  
DB 2664 CTTCTCTGCTCTGCGCCCTCTGCGCCCTCTGCTGCTGCTCTGCTGCTGCTGCT 2605  
QY 530 CATCCCGCTGCTGCGGAGCATGTACGCGGCTGCTGCGGCGGACGCTCATGCTCAGGT 589  
DB 2604 CTTGCGCCCTCTGCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2545  
QY 590 CGTCTGCTCTCAGTGTATCATGTGTACAGCTCATGCTCTTCTCTTCAGTACCGGCG 649  
DB 2544 CTTGCGCCCTCTGCGCCCTCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2485  
QY 650 CGCGCGCGCTGCTGCTTCTGACAGTTCGCGGCGGCGGCGGCTCCATGCTCTCTT 709  
DB 2484 CTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2427

QY 710 CGCGCTCGACTCCGACGTCTCTGCTCCCGAGGGGGAGCTCGAGCTCGAGGCCGAGCC 769  
DB 2426 CTGCGCCCTCTCTGCTGCTGCTGCGCCCTCTCTGCTGCTGCTGCTGCTGCTGCTG 2367  
QY 770 CGACGGCTCGCGGCGGCGGCGGCGGCTCTCTCTCCCGGCGGAGCGCGGGGGGTGG 829  
DB 2366 CTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2307  
QY 830 CGTCACCGTCCGCAAGTCCACAGCTCGGCTCCGAGGGCGGCTGCTC 877  
DB 2306 CCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 2259

Search completed: March 3, 2004, 06:21:36  
Job time : 220 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 3, 2004, 04:23:14 ; Search time 906 Seconds  
(without alignments)  
11025.470 Million cell updates/sec

Title: US-10-030-884-13

Perfect score: 2769

Sequence: 1 ccacgcctcgctgagcct.....aaaaaaaaaaaaaaaaaaag 2769

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2353733 seqs, 180373377 residues

Total number of hits satisfying chosen parameters: 4707466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgm2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgm2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgm2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgm2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgm2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgm2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgm2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgm2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgm2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgm2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgm2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgm2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 13: /cgm2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgm2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgm2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgm2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 17: /cgm2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 18: /cgm2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	836.8	30.2	1926	15	US-10-260-238-565
2	698	25.2	1759	15	US-10-260-238-30
3	681.2	24.6	2276	12	US-10-425-114-1349
4	612	22.1	2415	12	US-10-425-114-16771
5	586.4	21.2	651	15	US-10-260-238-5337
6	451.2	16.3	2346	12	US-10-424-599-134028
7	450	16.3	1273	12	US-10-425-114-15739
8	427.8	15.4	1313	12	US-10-425-114-11838
9	426.2	15.4	1334	12	US-10-425-114-2319
10	426.2	15.4	1199	12	US-10-425-114-29155
11	425.6	15.4	1321	12	US-10-425-114-33502
12	412.8	14.9	1091	12	US-10-425-114-3558
13	370	13.4	702	15	US-10-260-238-5464
14	347.2	12.5	2222	9	US-09-887-576-628
15	319.8	11.5	3020	12	US-10-424-599-45725

Sequence 848, App  
Sequence 848, App  
Sequence 4000, Ap  
Sequence 59026, A  
Sequence 847, App  
Sequence 847, App  
Sequence 104351,  
Sequence 134029,  
Sequence 86575, A  
Sequence 31117, A  
Sequence 134033,  
Sequence 29226, A  
Sequence 10077, A  
Sequence 7737, Ap  
Sequence 45726, A  
Sequence 1305, Ap  
Sequence 1305, Ap  
Sequence 45727, A  
Sequence 59027, A  
Sequence 71900, A  
Sequence 62375, A  
Sequence 5138, Ap  
Sequence 33472, A  
Sequence 13089, A  
Sequence 18686, A  
Sequence 52693, A  
Sequence 30762, A  
Sequence 2894, Ap  
Sequence 15102, A  
Sequence 1, Appli

#### ALIGNMENTS

#### RESULT 1

US-10-260-238-565  
Sequence 565, Application US/10260238  
Publication No. US20040016025A1  
GENERAL INFORMATION:  
APPLICANT: Budworth, Paul R.  
APPLICANT: Moughamer, Todd G.  
APPLICANT: Briggs, Steven P.  
APPLICANT: Cooper, Bret  
APPLICANT: Glazebrook, Jane  
APPLICANT: Goff, Stephen A.  
APPLICANT: Katagiri, Fumiyaki  
APPLICANT: Kreps, Joel  
APPLICANT: Provart, Nicholas  
APPLICANT: Ricke, Darrell  
APPLICANT: Zhu, Tong  
TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
FILE REFERENCE: 60111-NP  
CURRENT APPLICATION NUMBER: US/10/260,238  
CURRENT FILING DATE: 2002-09-26  
PRIOR APPLICATION NUMBER: US 60/325,448  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/325,277  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/370,620  
PRIOR FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 6077  
SEQ ID NO 565  
LENGTH: 1926  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: N region  
LOCATION: (261)..(261)  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N region  
LOCATION: (291)..(291)



PRIOR APPLICATION NUMBER: US 60/325,277  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: US 60/370,620  
PRIOR FILING DATE: 2002-04-04  
NUMBER OF SEQ ID NOS: 6077  
SEQ ID NO 30  
LENGTH: 1759  
TYPE: DNA  
ORGANISM: Oryza sativa  
FEATURE:  
NAME/KEY: N region  
LOCATION: (1584)  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N region  
LOCATION: (1587)  
OTHER INFORMATION: n = any nucleotide  
FEATURE:  
NAME/KEY: N region  
LOCATION: (1602)  
OTHER INFORMATION: n = any nucleotide  
US-10-260-238-30

Query Match 25.2%; Score 698; DB 15; Length 1759;  
Best Local Similarity 68.4%; Pred. No. 2.1e-160;  
Matches 1174; Conservative 0; Mismatches 438; Indels 105; Gaps 11;  
Qy 171 ATGATACCGCGCTGGACCTTACAGTGTCTGACGGCGGTGGTGGCGCTGTACGTGGCC 230  
Db 1 ATGATAACGGGGGGGACTTCTACACGTGATGACGGCGATGGTGGCGCTGTACGTGGCG 60  
Qy 231 ATGACGCTGGGTACGGCTCGCTCGGTGGTGGCGCATCTTACGCGCGACCAAGTGTCC 290  
Db 61 ATGATCTGGGTACGGGTGGTGGCGCATCTTACGCGCGACCAAGTGTCC 120  
Qy 291 GGGATCAACCGCTTGTGGCGCTCTTCCGCGTGGCGCTCTCTCTCTCTCTCTCTCTCTCT 350  
Db 121 GGGATCAACCGCTTGTGGCGCTCTTCCGCGTGGCGCTCTCTCTCTCTCTCTCTCTCTCT 180  
Qy 351 ACCAAGACCCCTTGGCGCATGACCTGCGTGGCGCGCATCTTACGCGCGACAGCTGCAAGGTG 410  
Db 181 ACCAAGACCCCTTACCATGACCTTGGCTTATTCGCGCGCGACACCTTGCAGAGCTC 240  
Qy 411 GCGGTCTCTGGCGCTGTGGCGCTGGCTTCCGCGCGCTCTCTCTCTCTCTCTCTCTCTCT 470  
Db 241 ATGCTCTCTGGCGCT 291  
Qy 471 CTCGACTGGAGATCACGCT 530  
Db 292 CTCGAGTGGACATCACCT 351  
Qy 531 ATCCCGCTGTGCGAGGATGTAAGGCGCTGTGCGCGCGCGCTCATGGTCCAGGTC 590  
Db 352 ATCCCGCTGTGAGGGGATGTAAGGCGAGTTCT--CCGCTAGCTCATGGTGCAGATC 408  
Qy 591 GTGCTCTCAGTGCATCATCTGTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 650  
Db 409 GTGCTCTCAGTGCATCATCTGTTACAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 468  
Qy 651 GCGCGCGCGCTGTCTCTGACAGTTTCCCGACGCGCGCGCGCGCTCATGCTCTCTCTCT 710  
Db 469 GCCAGATCTCTATCACCAGAGATTTCCCGACAGCTCTCTCTCTCTCTCTCTCTCTCTCT 525  
Qy 711 CCGCTGCACTCCAGTCT 770  
Db 526 GTGCTGGACGCGAGCTGTCT 585  
Qy 771 GACGGGTGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 830  
Db 586 GA-----GGTGAAGGAGGACCGGCAAGATAC 612  
Qy 831 GTCAACGCTGCGAAGTCCACAGCTCGGCGCTCCGAGGCGCGCTCTCTCTCTCTCTCTCT 887

Db 613 GTCAACGCTGCGCGCTCCAAAGCGGTGCGCTCCGAGCTTACTCCGCGGCTCCATGGG 672  
Qy 888 TCCAGACCATGACAGCCCGCTGTGTCCAACTCTCCGCGTGGAGATCTACTCGTGGAG 947  
Db 673 TTCTCCAGCACCAAGCGCGCGCGAGCAACTCTCAACAGCGGAGATCTACTCGTGGAG 732  
Qy 948 TCGTGGCGCAACCCACAGCGCGCGGTTCAGCTTCAACAGCGGAGATCTCTTCAACATC 1007  
Db 733 TCGTGGCGCAACCCACAGCGCGCGGTTCAGCTTCAACAGCGGAGATCTCTTCAACATC 792  
Qy 1008 GTGCGCGCGCGCGCGCGCGAGAGCGCGCGGAGCGCGCGGAGCGAGAGAGAGGCGC 1063  
Db 793 GTGCGCGCGCGCGCGCGCGAGCGCGCGGTTCGCGGTGCGCAACCGCGCGCAAG 852  
Qy 1064 -----ATGCGCGCGCGCGCGCGGAGAGCGCGCGGAGCGCGCGGAGCGCGCG 1107  
Db 853 CCGCGCGCGTCCAACTACAGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 912  
Qy 1108 TCGCGCGTCCG 1167  
Db 913 CCG 972  
Qy 1168 CCGGTGCGAGCGCG-----CCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1218  
Db 973 CCAAGCGCGAGAGCG 1032  
Qy 1219 ATGCGCGAGCTCTCCG 1278  
Db 1033 TGTCCAGCGCTTTCG 1092  
Qy 1279 CGAGCAGGAGAGAAAGCG 1338  
Db 1093 GCGTGGAGGAGCG 1152  
Qy 1339 TGGGTGCG 1398  
Db 1153 CCG 1194  
Qy 1399 CCGTGGCGAGTCCG 1458  
Db 1195 CGAGCGCGAGTCCG 1254  
Qy 1459 AGCTGATCGGAGCG 1518  
Db 1255 AGCTGATCGGAGCG 1314  
Qy 1519 CTT---ACAGGTGGCG 1575  
Db 1315 GCTTCCGCGAGGTGGAATTCGAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1374  
Qy 1576 CGAGCGCGGTCTCGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1634  
Db 1375 CGAGCGCGGTCTCGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1434  
Qy 1635 AGGATCATCGGTGCGGGAACAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1694  
Db 1435 CGGATCATCGGTGCGGGAACAGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1494  
Qy 1695 GCGCGCGCGGTCTATGCG 1754  
Db 1495 GCGCGCGCGGTCTATGCG 1554  
Qy 1755 ATCGCGCATCGT---CCAGGCTGCTCTGCGTCAAGGAGTCTGCGCGGTTCGTCGCAAG 1811  
Db 1555 GTCGCGCATGTGAGGAGGAGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1614  
Qy 1812 GAGTACGGGTTCATCCCGAGATCTCTGAGCAGCGGT 1848  
Db 1615 GAGTACAGGTGTCACCG 1651

RESULT 3  
US-10-425-114-1349



QY 1778 GCCTCAGGGATCGCCGTTGCTGTTCCGCAAGAGTACGGCTTCATCCCGACATCCT 1837  
DB 1857 TCCCAAGGCATCGFCCCTTCTGTTGTTTSCCAAGAGTACAACTGCGCATCTCAAACT 1916  
QY 1838 GAGCACAGCG 1847  
DB 1917 TAGCACAGCG 1926

## RESULT 4

US-10-425-114-16771  
; Sequence 16771, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jirongdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 16771  
; LENGTH: 2415  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3067-002-C10\_FLI  
US-10-425-114-16771

Query Match 22.1%; Score 612; DB 12; Length 2415;  
Best Local Similarity 63.4%; Pred. No. 2.8e-139;  
Matches 1125; Conservative 0; Mismatches 540; Indels 109; Gaps 8;  
QY 166 CGGCATGATCACCGCGTGGACCTTCTACAGTGTCTCAACGCGGTGACGCGGTGTCGCTGACG 225  
DB 184 CCGAGATGATATCGTGAACATGATATACAGGTGCTCAGCGCATGGTGCCTGTACG 243  
QY 226 TGGCATGACGTGCGTACCGCTCCGTCGCTGCGGCGATCTTACGCGGACCACT 285  
DB 244 TGGCATGTTCTGCGGTACGCGTCCGTCGCTGCGGCGATCTTACGCGGACCACT 303  
QY 286 GCTCCGGATCAACCGCTTCTGCGGCTCTTCCGCGTCCGCTCTCTCTCTTCACTCA 345  
DB 304 GCTCCGGATCAACCGCTTCTGCGGCTCTTCCGCGTCCGCTCTCTCTCTTCACTCA 363  
QY 346 TCTCCACCAACGACCCCTTCCGCGATGACCTGCGCTTCTGCGGCGGACGCGTGCAGA 405  
DB 364 TCTCCACCAACGACCCCTTCCGCGATGACCTGCGCTTCTGCGGCGGACGCGTGCAGA 423  
QY 406 AGTGGCGCTCTCGCGTCTGCGGCTGCGGCTCCGCGGCTCTCTCTCTCTCTCTCTCTCT 465  
DB 424 AGTGGCTCTCTCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 483  
QY 466 TCGGCTCGACTGAGATACGCT 525  
DB 484 CCGCGCTGGACTGGTCCATACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 543  
QY 526 TGGGATCCCGCTGCGGAGCATGTAGGCGGCTGCTGCGCGGCGCGTCTCATGTTGTC 585  
DB 544 TGGGATCCCGCTGCTGATGCGCATGTACGG--TCCCTACGCGGCTCTCTCTCTCTCTCTCT 600  
QY 586 AGTGGCT 645  
DB 601 AGTGGTGTCT 660  
QY 646 GCGCGCGCGCGGCT 705

DB 661 GCGCGCGCGGACGCTCATCGCGGACCACTTCCCGGA---CACCGCGCGGCGCATCGCGT 717  
QY 706 CTTTCGCTCGACTCCGACGCTGCTCTGCTCGCAGAGGGGAGCTTCGAGCTCGAGGCG 765  
DB 718 CGTGCACCTGACCCGCGACGTGGTGTCTGCTGAGGCGGCGCGCGCGAGACGAGGCGG 777  
QY 766 AGCCCGACGCGCTCGCGCGCGCGCGCGCTCTCTCTCCCGCGCGGCGAGCGCGCGCGG 825  
DB 778 AGTGGCGGAGGACGCGCGCTCCAGTCAAGTGGCGCGCTCTCTCTCTCTCTCTCTCTCT 837  
QY 826 TGGCGCTCAAGTGGCGCAAGTCAACAGCTCGCGCTCCGAGCGCGCGCTCTCTCTCTCTCT 885  
DB 838 CGTGTCTGATGCGAGCGCGCGCGCTCAACCTGACGCGCGCGAGATATATCTATGAG 897  
QY 886 ACTCCAGACCATGACGCGCGCTGCTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 945  
DB 898 CTGCTCGCGGACGACGCGCGCGCTCAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 957  
QY 946 AGTGTGCGGCAACCCACCGCGCGCGGTTCAGCTTCAACACGCGGACTTCTTCAACA 1005  
DB 958 GGTGACG 1013  
QY 1006 TCGTCCG 1065  
DB 1014 CCGCGAGCTGATCTGATGCT 1073  
QY 1066 GCG 1125  
DB 1074 ACGAGCGCTCGCGCTCG 1132  
QY 1126 GGAAGGACCTGCACATGCT 1185  
DB 1133 --AAGAGCTCCACATGTTCTGCTGAGCTTCGCGCGCTCTCTCTCTCTCTCTCTCTCT 1190  
QY 1186 CCGTGCACGCTTCT 1237  
DB 1191 GCCTGCGCGCTTCAACGCTGCTGCGCGCGCGCTCAACGCTCGCGCGCGCGCGCGCGCG 1250  
QY 1238 ----- 1237  
DB 1251 GCATGTTCT 1310  
QY 1238 -----AGGAGCGCGCGCGCT 1284  
DB 1311 ACGGTGACGACGACGACG 1370  
QY 1285 GGAAGAAAGCGGCG 1338  
DB 1371 AGACGCGCGCGGAGCG 1430  
QY 1339 TGGGGTCCA---ACTCGAGCGCGCGCGCTGTACCCCAAGGACGACGCGCGCGCGCGCGCG 1389  
DB 1431 TGGGGTCCAGGTCTCTGACGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1490  
QY 1390 GCG 1449  
DB 1491 GCTCG 1550  
QY 1450 TGTGAGGAGCTGATCCGGAACCCCAACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1509  
DB 1551 TGTGCGGAGCTGATCCGCAACCCCAACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1610  
QY 1510 CCTGTGTTCTTACAGTGGGCGATCGAGATGCCAGGATCATCGCGCGCGCGCGCGCGCGCG 1569  
DB 1611 CCTGTGCGGTTCGCTGGCGATCT 1670  
QY 1570 TCTGTGCGAGCGCGGTCTCGGGATGCGCATGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1629  
DB 1671 TCT 1730  
QY 1630 AGCGGAGGATCATCGCGCTCGGGAAACAAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1689  
DB 1731 AGCGGAGCTCATCGCTGCGGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1790

Qy 1690 TCGAGGCCCCGGGTCATGGCGCGCGCTCCATCGCCCTCGGTTCGGCGCGGTCCTCC 1749  
Db 1791 TCGCGGCGCTGCGGTCATGGCGCGCGCTCCATCGCCCTCGGTTCGGCGCGGTCCTCC 1850  
Qy 1750 TCCACATCCGCTGTCAGGTCGCTCGCTCAGGGATCGTCCGTTCTCGGTCGCA 1809  
Db 1851 TCGGATCGCCATGTCAGGCGGCTCTACCGAAGGATCGTCCGCTTCGTTTCGCA 1910  
Qy 1810 AGGATACGCGCTTCATCCCGACATCCTGAGCAC 1843  
Db 1911 AAGATACAGCTCCACCTGCTGAGCAC 1944  
RESULT 5  
US-10-260-238-5337  
; Sequence 5337, Application US/10260238  
; Publication No. US20040016025A1  
; GENERAL INFORMATION:  
; APPLICANT: Budworth, Paul R.  
; APPLICANT: Moughamer, Todd G.  
; APPLICANT: Briggs, Steven P.  
; APPLICANT: Cooper, Bret  
; APPLICANT: Glazebrook, Jane  
; APPLICANT: Goff, Stephen A.  
; APPLICANT: Katagiri, Fumiyaki  
; APPLICANT: Kreps, Joel  
; APPLICANT: Provart, Nicholas  
; APPLICANT: Ricke, Darrell  
; APPLICANT: Zhu, Tong  
; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION  
; FILE REFERENCE: 60111-NP  
; CURRENT APPLICATION NUMBER: US/10/260,238  
; CURRENT FILING DATE: 2002-09-26  
; PRIOR APPLICATION NUMBER: US 60/325,448  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: US 60/325,277  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: US 60/370,620  
; PRIOR FILING DATE: 2002-04-04  
; NUMBER OF SEQ ID NOS: 6077  
; SEQ ID NO 5337  
; LENGTH: 651  
; TYPE: DNA  
; ORGANISM: Zea mays  
US-10-260-238-5337  
Query Match 21.2%; Score 586.4; DB 15; Length 651;  
Best Local Similarity 99.7%; Pred. No. 2.8e-133;  
Matches 598; Conservative 0; Mismatches 1; Indels 1; Gaps 1;  
Qy 1238 AGGAGCCAGGCTTACGACGAGTACGGGCGCGACGACTACGACGAGCAGGACGAG 1297  
Db 52 AGGAGCCAGGCTTACGACGAGTACGGGCGCGACGACTACGACGAGCAGGACGAG 111  
Qy 1298 GAGGCGGCGCGGACGAGGCGGCGCGCTCGAGCTGGGTCCAACTCGACGCG 1357  
Db 112 GAGGCGGCGCGGACGAGGCGGCGCGCTCGAGCTGGGTCCAACTCGACGCG 171  
Qy 1358 GCAGCTGTACCCCAAGGACGACGCGGAGGAGGCGGCGCGCTCGAGTCCGCGCG 1417  
Db 172 GCAGCTGTACCCCAAGGACGACGCGGAGGAGGCGGCGCGCTCGAGTCCGCGCG 231  
Qy 1418 GAGGCTATGACGCGGCTCATCTTCATGTTGTGGAGGAAGCTGATCGGAAACCCAA 1477  
Db 232 GAGGCTATGACGCGGCTCATCTTCATGTTGTGGAGGAAGCTGATCGGAAACCCAA 291  
Qy 1478 CACCTACTCCAGCCTCATCGGCGCTCGTCTGGTCCCTCGTCTCTACAGTGGGCGATCGA 1537  
Db 292 CACCTACTCCAGCCTCATCGGCGCTCGTCTGGTCCCTCGTCTCTACAGTGGGCGATCGA 351  
Qy 1538 GATCCAGCAGTATCATCGCCCGGTCGATTCGATCCTGTCCGACGCGGCTTCGGATGCG 1597

Db 352 GATGCCAGCGATCATCGCCCGGTCGATTCGATCCTCTCGGACGCGGTCCTCGGATGCC 411  
Qy 1598 CATGTTCAGCTAGGCTGTTCATGCGCTGACGCGAGGATCATCGCTGCGGGAACAA 1657  
Db 412 CATGTTCAGCTAGGCTGTTCATGCGCTGACGCGAGGATCATCGCTGCGGGAACAA 471  
Qy 1658 GCTGGCGGCGCATCGCATGCGGCTCGGTTCTCGAGGCGCGG-GGTTCATGGCGCGCG 1716  
Db 472 GCTGGCGGCGCATCGCATGCGGCTCGGTTCTCGAGGCGCGGCGGTCATGGCGCGCG 531  
Qy 1717 CTTCCATCGCGCTGCTCGCGGCGTCTCTCCATCGCATGCTCCAGGCTGCTC 1776  
Db 532 CTTCCATCGCGCTGCTCGCGGCGTCTCTCTCCATCGCATGCTCCAGGCTGCTC 591  
Qy 1777 TGCCTCAGGGATCGTCCGCTTCGTTCCGCAAGGAGTACGCGTTTCATCCCGACATCC 1836  
Db 592 TGCCTCAGGGATCGTCCGCTTCGTTCCGCAAGGAGTACGCGTTTCATCCCGACATCC 651  
RESULT 6  
US-10-424-599-134028  
; Sequence 134028, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J  
; APPLICANT: Kovalic, David K  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 134028  
; LENGTH: 2346  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_92032C.1  
US-10-424-599-134028  
Query Match 16.3%; Score 451.2; DB 12; Length 2346;  
Best Local Similarity 58.1%; Pred. No. 6.3e-100;  
Matches 1020; Conservative 0; Mismatches 643; Indels 93; Gaps 9;  
Qy 164 CGCGCCATGATCACCGCGTGGACCTCTACACGCTCTGACGCGGTCGTGGCGCTGTA 223  
Db 141 CACGACAATGATAACGTTGTTGGACTTGTACACGCTTCTGACAGCGGTGGTGGCGCTTA 200  
Qy 224 CGTGGCATGACGCTGGGTACGCTCGCTCGGTGGCGATCTTTCACGCGGACCA 283  
Db 201 CGTGGCATGATCTCGCTACGCTCGGTGAGTGGTGAAGATCTTACCGCGGACCA 260  
Qy 284 GTGCTCGGGATCAACCGCTTCGTGGCGCTTTCGCGTGGCGCTCTCTCTTCCTTCACTT 343  
Db 261 ATGCTCGGGAATAACCGTTTCGTGGCACCTTCGAGTGGCGCTCTCTCTCTTCCACTT 320  
Qy 344 CATCTCCACCAACGACCGCTTCGCGATGACCTGCTGCTGGCGCGCGGACAGCTGCA 403  
Db 321 CATCTCCACCAACGACCGCTTCGCGATGACCTGCTGCTGGCGCGGACAGCTGCA 380  
Qy 404 GAAGTGGCGCTCTCGCGTTCGTGGCGTGGCGTGGCGCTTCGCGGCGCTCTCTCTCCCGCGCG 463  
Db 381 AAAGCCATAGTCTCTCGAGTGTCTTGGTGGTCCAGAACAAAGCTC-----AAG 431  
Qy 464 GCTCGGCTCGACGAGCATCAGCTCTTCTCTCTCTCCAGCTTCCCAACAGCTCGT 523  
Db 432 AGGCTCATTGAATGGTCCATAACACTCTTCTCTCTCACTCTCCCAACACTTGGT 491  
Qy 524 CATGGGCACTCCCGCTGCTGCGAGGATGATACGCGCGCTGCTGCGCGGACAGCTCATGGT 583  
Db 492 GATGGTATCCCGTGTGCTGAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 548

QY 584 CAGGTGCTGCTCTCCAGTGCATCATCTGTGTAACGCTCATGCTCTCTCTTCGAGTA 643  
Db TCAGATTGTGCTGCTTCACTGATCATCTGTATACCTTATGTTGTTCTTTTCGAGTA 608  
QY 644 CCGCGCGCGCGCGCTGCTCTGACACAGTTCCTCGACGGCGCGCGCTCCATCGT 703  
Db TAGGGTGGAGGCTTTTAAATAGTGGACAGTTTCTTGA----TACAGCGGTTGATAT 665  
QY 704 CTCCTTCGCTGCACTCCGACGTGCTCTGCTCGCCAGGGGGACGTGAGCTCGAGGC 763  
Db CTCCTTCAAGTTGATTCCGATATTATTTCTGATGGGAAGAGCGCTTCAGACGGA 725  
QY 764 CGAGCCGACGCG-----GTGCGCGCGCGCGC 791  
Db GCGGAGGTTGTTGATGACGGGAAGCTTCATGTGACGGTGAAGTCCGAGTTCCGG 785  
QY 792 GCCGTCTCTCCCGCGCGGGAGCGCGGGGGTGGCGGTCACTCCGACGACGACCGCG 847  
Db GTCGAGATTCTCTAGCGCTCTCATGGCCCCAACCTCGGTGCTGTTGACTCCGAGGCC 845  
QY 848 CACGAGCTGCGCTCCGAGGCGGGTGTCTGCACTCGCATCCGACACCATGACGCGCG 907  
Db TTCGAATTTGACCAATGCGGAGATTTACTCGCTCAGAGCTTCGAGGAACCGACTCCGAG 905  
QY 908 TGTGTCCAACTCTCCGCGCTGGAGATCTACTCGCTGCAGTCTGCGGCAACCCACCCC 967  
Db AGGCTCGAGTTTAAACACACGGAATTTTACTTCCATGTTGATGTTGAACACACAA 965  
QY 968 GCGGGGTCCAGCTTCAACACGCGGACTTTTCAACATGCTGCGCGCGCGCGCAAGG 1027  
Db CAACAAACAAACAGTGTAGTCCGAGGAGTAAATTTTGGAGGCTTTGATGAAGA 1025  
QY 1028 AGCGGAGGAGCGCGGGGACGAGG-----AGAGGGCGCATGCGGCGC 1073  
Db AAGTGTGCGGTGTGAGGGTGAATGTGTGCTGTGTGCTGCTGCTGCTGCTGCTGCTG 1085  
QY 1074 GCGCGCGGAGGACACT-----CGCGGAGCGCGAGCGCTGCGCGTGC CGGCGCAA 1123  
Db GAATCGCGGGATTTTCTCGGTGGCGCGGAGAGGAGGTGTGAGAGTGTGTGTGG 1145  
QY 1124 GAGGAAGACCTGCATGCTGCTGAGCTCCAGCGCTCGCGCGCTGCTGCGAGCGCGC 1183  
Db AGGAGAGATCTCCACATGTTTGTGTGAGTTTCAAGTCTTCCGCGGTGCTGCGAGGTGG 1205  
QY 1184 CGCGTGCACGTCTTTCGCGCGCGCGCTGACCATGCGGACGCTCTCGCCAAAGAGC 1243  
Db AATCCATGCTTTCAGAGTGTGTGTGGGATTATGGGAGTCAACAGCTTCTCTGTGTGG 1265  
QY 1244 CCAGGCC-----TAGCAGAGTACGGGCGCGGACACTACAGCAGCAGACGAA 1291  
Db GGTGGCTCACCAAGAAATATGATGATTTTGGTACAGATGATTTAGCTTCGGGAACAG 1325  
QY 1292 GAAACGGAGCGCGCGCGGCAAGGGCGCGCGAGCTGTCGAAGCTGGGGTCCAACTC 1351  
Db AACCTTCTAATGGGTTGACAGGAAGGCGAGTGTCTTCAAGCTTGGCTCGAGTTC 1385  
QY 1352 GACGCGGAGCTGTATACCCCAAGGACGACGCGGAGGGCGCGCGGTGGCGATGCC 1411  
Db CACAGCTGAGCTTCGCCCTTAAAGCTCAAGGTGAAGCCAAACCTTAC-----TTCCATGCC 1439  
QY 1412 GCGCGGAGCGGTGATGACGCGCTCATCTCATATGTTGTGAGGAAGCTGATCCGAA 1471  
Db ACCAAAGAGTGTATGACAGGCTCATTTTGTATGTTTGGAGGAAGCTGATAGAA 1499  
QY 1472 CCCCACACCTACTCCAGCTCATCGGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1531  
Db CCCCACACATATTCAGCTCTTTGGTCTCACTTGGTCTTTGATCTCATTTCAATGAA 1559  
QY 1532 CATCGAGTCCAGGATCATCGCGGCTGATTTTCGATCTGTGCGAGCGGCTCTCGG 1591  
Db TGTGTTGTCGAGCAATTTGTTGCTAAATCGATATCAATTTTATGATGCTGTGCTTGG 1619

QY 1592 GATGCCCATGTTTACGCTAGGCTGTTTATGCGGCTGACGCGGAGATCATCGCGTGGG 1651  
Db GATGCCAATGTTTACGCTTGGCTTATTATGCGCATGACGCGCAAGATTTTGCATGTGG 1679  
QY 1652 GAACAGCTGGCGGCGCATCGCATGGGCGTCCGTTGCTGCGAGGCGCGCGGTGATGCG 1711  
Db AACTCGGTTGCTTCTTTGCTATGCGAGTTCGTTTCTTACTGCTCGCATGATGCG 1739  
QY 1712 CGCGGCTTCATCGCGTGGTGGTGGCGGCGTCTCTCCACATCGCATCGCATCGCG 1771  
Db TGTGCTTCAATCGTTGTTAGGCTCAGGGAGTTCTATTGCGCATGCTGCTTGTACAGG 1799  
QY 1772 TGTCTGCTCAGGGGATCGTCCGTTGCTGTTCCGCAAGGAGTACGGGTTTCATCCCGA 1831  
Db TGTCTGCTCAGGGATGTTGCTTCTTGTGTTGCTAAGGAATACAGGTTTCATCTGA 1859  
QY 1832 CATCTGAGCACAGCG 1847  
Db CATACTAAGCACCGGG 1875

RESULT 7  
US-10-425-114-15739  
; Sequence 15739, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaka, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 15739  
; LENGTH: 1273  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB3060-064-E7\_FLI  
US-10-425-114-15739

Query Match 16.3%; Score 450; DB 12; Length 1273;  
Best Local Similarity 66.7%; Pred. No. 9.4e-100;  
Matches 743; Conservative 0; Mismatches 320; Indels 51; Gaps 5;

QY 171 ATGATCACCGCGTGAACCTTACACGCTGCTGACGCGGCTGCTGCGCTGTACGTGGCC 230  
Db 188 ATGATTCGCGCGCGGACTTGTACACGCTGATGACGCGCGTGTGCGCTGTACGTGGG 247  
QY 231 ATGACGCTCGCTGACGCTCCGCTCCGCTGGTGGCGCATCTTCAACGCGGACCATGCTCC 290  
Db 248 ATGATCCTGGCTGACGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 307  
QY 291 GGGATCAACGCTTGTGCGGCTCTTCGCGCTGCGGCTCTCTCTCTTCCACTTCACTCC 350  
Db 308 GGGATCAACGCTTGTGCGGCTCTTCGCGCTGCGGCTCTCTCTCTTCCACTTCACTCC 367  
QY 351 ACCAAGACCCCTTCCCATGAACCTTCCGCTGCTGCGCGCGCGGACACGCTGACGAGGTC 410  
Db 368 ACCAACAACCCCTACACCATGAACCTGCGCTTTCATCGCGCGCGGACACGCTGACGAGCTC 427  
QY 411 GCGCTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 467  
Db 428 ATGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 487  
QY 468 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527  
Db 488 GACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547

```
QY 528 GGCATCCCGCTGTGCGAGCATGTACGGCGGGTGGTGGCGCGGACGCTCATGCTCCAG 587
Db 548 GGCATCCCGCTGTGCGAGCATGTACGGCGGGTGGTGGCGCGGACGCTCATGCTCCAG 604
QY 588 GTGCTGCTCTCGAGTGCATCATCTGGTACAGCTCATGCTCTTCTCTTCCAGTACCGC 647
Db 605 ATCGTGTGCTCCAGTGCATCATCTGGTACAGCTCATGCTCTTCTCTTCCAGTACCGC 664
QY 648 GCGCGCGCGCGCTCGCTCTCCAGCAGTTCGCCGAGCGCGCGCGCTCCATGCTCTCC 707
Db 665 GCGCGCGCGCTCATCTCCGAGCAGTTCGCCGAGCGCGCGCGCTCCATGCTCTCC 724
QY 708 TTCGCGTGCAGTTCGAGCTGTGTCTGCTCGCGCAGGGGGGAGCTCGAGCTCGAGCCGAG 767
Db 725 ATCGCGCTGCAGCCGAGCTGTGTCTCCCTCGTC----- 757
QY 768 CCGCAGCGGGTGCCTCGCGCGCGCGCGCTCTCTCTCCGCGCGGGGAGCGCGCGGGGTG 827
Db 758 GACCAAGCGCGGGGAGCGCATCGAGAGCGAGCGCGAGTCAAGGAGGAGCGGAGGATA 817
QY 828 CCGCTCAAGTGGGGAAGTTCACAGCTCGCGTTCGAGCGCGCGCTGCTCGCAC----- 881
Db 818 CAGCTCAGCTGCGCGCTCCAGCGCTGCGCTCGAGCTGCTCTACTCGCGCGCGCTCC 877
QY 882 ---TCGCACTCCAGACCATGACAGCCCGTGTGTCCAACTCTCCGGGTGAGATCTAC 938
Db 878 GTGGCGGGTCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 937
QY 939 TCCTGAGTGTGCGGCAACCCACCGCGCGGGTTCAGTTCACACCGCGGAGTTC 998
Db 938 TCCTGAGTGTGCGGCAACCCACCGCGCGGGATCCAGTTCACACCGCGGAGTTC 997
QY 999 TTCACATGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1058
Db 998 TACTCCATGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1057
QY 1059 GCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1109
Db 1058 ACGCCGCGCGCGTCCAACTACGAGGAGCGCGTCCAGCGCGCGCGCGCGCGCGCG 1117
QY 1110 GCGTGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1169
Db 1118 GTGGCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1177
QY 1170 GTCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1229
Db 1178 AAGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1237
QY 1230 CTCGCCAAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1263
Db 1238 TTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1271
```

## RESULT 8

```
US-10-425-114-11838
; Sequence 11838, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 11838
; LENGTH: 1313
```

```
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 701184321_FLI
US-10-425-114-11838

Query Match      15.4%; Score 427.8; DB 12; Length 1313;
Best Local Similarity 72.0%; Pred. NO. 2.6e-84;
Matches 650; Conservative 0; Mismatches 202; Indels 51; Gaps 5;

QY 164 CCGCGCATGATCACCGCGCTGAGACCTCTACACGCTGAGCGCGGTGGTCCGCTGTA 223
Db 188 CTCACGATGATTACGGGAGCGGACTTCTACACGCTCATGACGCGGTGGTCCGCTGTA 247
QY 224 GTGGCATTACGCTGGCGTACCGCTCCGCTGCTGGTGGCGCATCTTACGCGCGGACCA 283
Db 248 CGTGGCATGATCTTGGCGTACCGGTGCGGTGGTGGCGCATCTTCTCGCGGACCA 307
QY 284 GTGCTCCGGGATCAACCGCTTCTGGCGCTCTTTCGCGTGGCGCTCTCTCTTCCTTCCACTT 343
Db 308 GTGCTCCGGGATCAACCGCTTCTGGCGCTCTTTCGCGTGGCGCTCTCTCTTCCTTCCACTT 367
QY 344 CATCTCCACCAACGACCGCTTTCGCCATGACCTGCGCTTCTCGCGCGCGGACGCTGCA 403
Db 368 CATCTCCACCAACGACCGCTTTCGCCATGACCTGCGCTTCTCGCGCGCGGACGCTGCA 427
QY 404 GAAGGTGGCGCTCTCTCGCGTGTGGCGTGGCGCTCTCGCGCGCTCTCTCTTCCTTCGCGCG 463
Db 428 GAAG-----CTCATGTGTGTGGCATGTCTCACGGGTGGAGCCACCTCAGCGCGCG 478
QY 464 GCTCGGCTCGACTGGAGCATCAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523
Db 479 GGGCAGCTCGAGTGGACCATCAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 538
QY 524 CATGGGATCCCGTGTGGAGCATGTACGGCGCTCTCGCGCGCGGACGCTCATGCT 583
Db 539 CATGGGATCCCGTGTGGAGCATGTACGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 595
QY 584 CCAGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 643
Db 596 GCAGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 655
QY 644 CCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 703
Db 656 CCGCGCGCGCGCGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 712
QY 704 CTCCTTCGCGTCTGACTCCGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 763
Db 713 CTCCTTCGCGTCTGACTCCGACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 772
QY 764 CGAGCCGAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823
Db 773 GGAGGCGA-----GGTCAAGGAGGAGCGGAG 799
QY 824 GGTGCGGTCAACGCTGGCGCAAGTCCACAGTCTGCGCTCCGA---GGCGCGGTGCTCGCA 880
Db 800 GATACAGTCAACGCTGGCGCGCTCCAGCGCTCGCGCTCCGACATCTACTCTCGCGCGCTC 859
QY 881 CTCGCACTCCAGACCATGAGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 940
Db 860 CATGGCTTCTTCCAGCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 919
QY 941 GCTGCGTCTGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1000
Db 920 GCTGCGTCTGCGCAACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 979
QY 1001 CAACATGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1060
Db 980 CTCATGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1039
QY 1061 CGC 1063
Db 1040 CGC 1042
```

## RESULT 9

US-10-425-114-2319  
; Sequence 2319, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 2319  
; SEQ ID NO 2319  
; LENGTH: 1334  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700209522\_FLI  
US-10-425-114-2319

Query Match 15.4%; Score 427.8; DB 12; Length 1334;  
Best Local Similarity 72.0%; Pred. No. 2.6e-94;  
Matches 650; Conservative 0; Mismatches 202; Indels 51; Gaps 5;  
QY 164 CGCGCCATGATACCGCGCTGGAGCTTACAGTGTGACGGCGGTGGTGGCGCTGTA 223  
DB 210 CTCACAGTATACCGGAGCGGACTTCTACACGCTCATGACGCGCGGTGGTGGCGCTGTA 269  
QY 224 CGTGCCATGACGTGGGCTACGCTCGCTCGGTGGGCGCATCTTACGCGCGACCA 283  
DB 270 CGTGCGATGATCTTGGCGGTACGGTGGTGGTGGGCGATCTTTCGCGCGACCA 329  
QY 284 GTGCTCCGGATCAACCGCTTGTGGCGCTTTGCGCGTGGCGCTCTCTCTTCCACTT 343  
DB 330 GTGCTCCGGATCAACCGCTTGTGGCGCTTTGCGCGTGGCGCTCTCTCTTCCACTT 389  
QY 344 CATCTCCACCAAGACCGCTTGGCATCAAGCTGGCTTCTGCGCGCGCGACGCTGCA 403  
DB 390 CATCTCCACCAAGACCGCTTGGCATCAAGCTGGCTTCTGCGCGCGCGACGCTGCA 449  
QY 404 GAAGTGGCGGCTCTGCGCTGCTGCGCTGCTGCGCGCGCTCTCTCTCTCTCTCTCT 463  
DB 450 GAAG-----CTCATGGTGTGGCATGCTCACGCGTGGAGCCACCTCAGCGCGG 500  
QY 464 GCTCGGGCTGACTGGAGATACGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523  
DB 501 GGGCAGCTGGAGTGGACCATACGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 560  
QY 524 CATGGGCATCCCGCTGCTCGAGGATGTACGGCGCTGCTGCGCGCGCGCTCATGGT 583  
DB 561 CATGGGCATCCCGCTGCTCGAGGATGTACGGCGCTTCT-----CGGCGAGCTTCT 617  
QY 584 CCAGTGTGTGCT 643  
DB 618 GCAGATGTGTGCT 677  
QY 644 CGCGCGCGCGCGCGCTGCTCTGACAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 703  
DB 678 CGCGCGCGCGCGCGATGCTCATACGAGAGATTTCCCGA-----CAACGCGCGCGCATCG 734  
QY 704 CTCCTTCCGCGTCTGATCCGAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 763  
DB 735 CTCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 794  
QY 764 CGAGCCGAGCGGTGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 823

DB 795 GGAGGCCGA-----GGTCAAGGAGGACGGCAG 821  
QY 824 GGTGCGCTTACCGTGGCGAAGTCCACAGCTCGGCTCCGA---GGCGCGCTGCTCGCA 880  
DB 822 GATACAGTACCGTGGCGCGCTTCAAGCTTCGCGCTCCGACATCTACTCGCGCGCTC 881  
QY 881 CTCGCACTCCAGACCATGACAGCCCGGTGTGTCAACCTCTCCGCGGTGGAGATCTACTC 940  
DB 882 CATGGCTTCTCCAGACCAACCGCGCGCGCGAGCAACCTGACCAACGCGGAGATCTACTC 941  
QY 941 GCTGAGTGTGCGCGCAACCGCGCGCGCGCGGTTCAGCTTCAACCAACGCGGACTTCTT 1000  
DB 942 GCTGAGTGTGCGCGCAACCGCGCGCGCGCGGTTCAGCTTCAACCAACGCGGACTTCTT 1001  
QY 1001 CAACATCTGTCGCGCGCGCGCGCAAGGAGGCGGCGGAGGCGCGGCGGAGGAGG 1060  
DB 1002 TCCCATGTGTCGCGCGCGCTCAACTTTCGCGCGCGCGCGCTTCGGCATCCGACCGG 1061  
QY 1061 CGC 1063  
DB 1062 CGC 1064

## RESULT 10

US-10-425-114-29155  
; Sequence 29155, Application US/10425114  
; Publication No. US2004003488A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(5313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 29155  
; LENGTH: 1199  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB83-006-E4\_FLI  
US-10-425-114-29155

Query Match 15.4%; Score 426.2; DB 12; Length 1199;  
Best Local Similarity 71.9%; Pred. No. 6.1e-94;  
Matches 649; Conservative 0; Mismatches 203; Indels 51; Gaps 5;  
QY 164 CGCGGCCATGATCACCGCTGGACCTCTACCACTGCTGAGCGGCGGTGGTGGCGCTGTA 223  
DB 77 CTCACAGTATACCGGAGCGGACTTCTTACCACTGATGACGGCGGTGGTGGCGCTGTA 136  
QY 224 CGTGCCATGACGTGGCGTACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 283  
DB 137 CGTGCGATGATCTCTGCGCTACGGGTGGTGGCGGTGGTGGCGCATCTTCTCGCGGACCA 196  
QY 284 GTGCTCGGGATCAACCGCTTGTGGCGCTCTTTCGCGGTGGCGCTCTCTCTCTCTCTCTCT 343  
DB 197 GTGCTCGGGATCAACCGCTTGTGGCGCTCTTTCGCGGTGGCGCTCTCTCTCTCTCTCTCT 256  
QY 344 CATCTCCACCAAGACCGCTTTCGCACTGAACTGGCGTCTCTGGCGCGCGACGAGTGA 403  
DB 257 CATCTCCACCAAGACCGCTTACACCTGAACTGGCGTCTCTTCATCGCGCGGACGAGTGA 316  
QY 404 GAAGTGGCGCTCTCGCGTGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCG 463  
DB 317 GAAG-----CTCATGGTGTGGCGCATGTCTACCGGTGGAGCACTCTAGCGCGCG 367  
QY 464 GCTCGGCTCGACTGGAGCATCAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523



; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

; FILE REFERENCE: 38-21(5313)B

; CURRENT APPLICATION NUMBER: US/10/425,114

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 73128

; SEQ ID NO 3558

; LENGTH: 1091

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: 700321691\_FLI

US-10-425-114-3558

Query Match 14.9%; Score 412.8; DB 12; Length 1091;  
Best Local Similarity 71.2%; Pred. No. 1.1e-90;  
Matches 638; Conservative 0; Mismatches 207; Indels 51; Gaps 5;

```
Qy 171 ATGATCACCGCGCTGACCTCTACACGTCGTGACGCGGTGTGCGCGCTGTAGTGGCC 230
Db 12 ATGATCACCGCGCGACTTCTACACGTCATGACGCGCATGGTGGCGTGTAGTGGCC 71

Qy 231 ATGAGCTGGCGTAGCGTTCGTCGCTGGTGGGCGATCTTCACGCGGACAGTGTCC 290
Db 72 ATGATCTCGGCTACGATCGTTCAGGTGGTGGCGCATCTTCACGCGGACCAAGTGTCC 131

Qy 291 GGGATCAACCGCTTCGTGGCGCTCTTCGCGTGGCGCTCTCTCTCTCTCTCTCTCTCT 350
Db 132 GGGATCAACCGCTTCGTGGCGCTCTTCGCGTGGCGCTCTCTCTCTCTCTCTCTCTCT 191

Qy 351 ACCAAGACCCCTTCGCCCATGAACCTGCGCTTCCTGGCGCGGACACAGCTGCGAAGTG 410
Db 192 ACCAAGACCCCTTCGCCCATGAACCTGCGCTTCCTGGCGCGGACACAGCTGCGAAGTG 251

Qy 411 GCCGTCTCGCGCTGCTGGCGCTGCGCTCGCGCGCTCTCTCTCTCTCTCTCTCTCTCT 470
Db 252 ATCGTCTCGCGCTGCT-----CACTGCATGAGACTACCTCTCCGCGGGGTGC 302

Qy 471 CTCGACTGAGCATACAGCTCTTCTCCCTCTCCAGCTTCCCAACAGCTCGTCAAGGCG 530
Db 303 CTCGAGTGGACATACAGCTCTTCTCCCTCTCGAGCTGCGCAACAGCTGATGGCG 362

Qy 531 ATCCGCTGCTCGAGGATGATGCGCGCGTCTCGCGCGCGCGCGCGCGCGCGCGCGCG 590
Db 363 ATCCGCTGCTCGAGGATGATGCGCGCGTCTCT-----CGGCGACCTCATGGTGCAGATC 419

Qy 591 GTCGTCTCTCCAGTGCATCATCTGTATACAGCTCATGCTCTTCTCTCTCTCTCTCTCT 650
Db 420 GTGGTCTCTCCAGTGCATCATCTGTATACAGCTCATGCTCTTCTCTCTCTCTCTCTCT 479

Qy 651 GCGCGCGGCTGCTCTCGACAGTTCCTCCGACGCGCGCGCGCGCGCGCGCGCGCGCG 710
Db 480 GCGAGGATCTCTATACCGAGCAGTTCCTCCGACGCGCGCGCGCGCGCGCGCGCGCG 536

Qy 711 CGCGTCGACTCCGAGTCTGCTCGCTCGCCAGGGGAGCGCTCGAGCTCGAGGCGGAGCC 770
Db 537 GTGGTGCACCCGACGCTGCTGCTGGAAGGGCGCAACGACGCGCATCGAGCGAGGCC 596

Qy 771 GACGCGCTCGCGCGCGCGCGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 830
Db 597 GA-----GGTGAAGGAGGACCGCAAGATACAC 623

Qy 831 GTCACCGTGCAGAGTCCACAGCTCGCGCTCCGAGGCGCGCTCGCACTC---GCAC 887
Db 624 GTCACCGTGCAGGCTCCAAAGCGTTCGCGCTCGGACATCTACTCCGCGCGTCCATGGG 683

Qy 888 TCCAGACCATGACGCCCGCTGTCTCAACCTCTCTCCGCGTGGAGATCTACTCGCTCAG 947
Db 684 TTCACAGCACACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 743

Qy 948 TCGTCGCGAACCACCCCGCGCGCGGCTGAGCTTCACACCGCGGAGCTCTTCAATC 1007
Db 744 TCGTCGAGGAACCCACCGCGCGCGGCTCCAGCTTCACACCGCGGAGCTCTTCTACTCCATG 803
```

```
Qy 1008 GTCCGCGCGCGCGCCCAAGGAGCGGAGCGGAGCGCGCGGAGCGAGAGGAGCGGC 1063
Db 804 GTCCGCGCGCGAGCTCCAACTTCGCCCGCGGAGCGGCTTCGGCGCTGCGACAGCGGC 859
```

RESULT 13

US-10-260-238-5464/c

; Sequence 5464, Application US/10260238

; Publication No. US2004016025A1

; GENERAL INFORMATION:

; APPLICANT: Moughamer, Paul R.

; APPLICANT: Briggs, Steven P.

; APPLICANT: Cooper, Bret

; APPLICANT: Glazebrook, Jane

; APPLICANT: Goff, Stephen A.

; APPLICANT: Katagiri, Fumiaki

; APPLICANT: Kreps, Joel

; APPLICANT: Provart, Nicholas

; APPLICANT: Ricke, Darrell

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: PROMOTERS FOR REGULATION OF PLANT EXPRESSION

; FILE REFERENCE: 60111-NP

; CURRENT APPLICATION NUMBER: US/10/260,238

; CURRENT FILING DATE: 2002-09-26

; PRIOR APPLICATION NUMBER: US 60/325,448

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US 60/325,277

; PRIOR FILING DATE: 2001-09-26

; PRIOR APPLICATION NUMBER: US 60/370,620

; PRIOR FILING DATE: 2002-04-04

; NUMBER OF SEQ ID NOS: 6077

; SEQ ID NO 5464

; LENGTH: 702

; TYPE: DNA

; ORGANISM: Zea mays

US-10-260-238-5464

Query Match 13.4%; Score 370; DB 15; Length 702;

Best Local Similarity 78.4%; Pred. No. 2.7e-80;

Matches 486; Conservative 0; Mismatches 120; Indels 14; Gaps 3;

```
Qy 171 ATGATCACCGCGCTGACCTCTACACGTCGTGACGCGGTGTGCGCGCTGTACGTGGCC 230
Db 651 ATGATTTACGGGACGAGCTTCTACACGTCATGACGCGCGTGTGCGCTGTACGTGGCG 592

Qy 231 ATGACGCTGGCGTAGGGCTCCGTCGCTGGTGGCGCATCTTCACGCGCGACCGAGTCTCC 290
Db 591 ATGATCTGGCTACGGGTGCGTGGTGGCGCATCTTCTCGCGGACCGAGTCTCC 532

Qy 291 GGGATCAACCGCTTCGTGGCGCTCTTGGCGTGGCGCTCTCTCTCTCTCTCTCTCTCTCT 350
Db 531 GGGATCAACCGCTTCGTGGCGCTCTTGGCGTGGCGCTCTCTCTCTCTCTCTCTCTCTCT 472

Qy 351 ACCAAGACCCCTTCGCCCATGAACCTGCGCTTCTCTGGCGCGCGACACGCTGCAGAGGTG 410
Db 471 ACCAAGACCCCTTCACCATGAACCTGCGTTCATCGCGCGCGACACGCTGCAGAGGTG 412

Qy 411 GCGGTCTCGCGCTGCTGGCGTGGCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 470
Db 411 ATGGTCTGGCGATGCT-----CACGCGTGGAGCCACCTCAGCGCGCGGCGAGCC 360

Qy 471 CTCGACTGAGCATACGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 530
Db 359 CTGAGTGAACCATACGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300

Qy 531 ATCCCGCTCTCGAGGCGATGTACGCGCGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 590
Db 299 ATCCCGCTCTCAAGGCGATGTACGCGCGCTCTCT---CGGCGAGCTCTCATGTTGCAGATC 243

Qy 591 GTCGTCTCCAGTGCATCATCTGCTACAGCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCT 650
Db 242 GTCGTCTCCAGTGCATCATCTGCTACAGCTCATGCTCTCTCTCTCTCTCTCTCTCTCTCT 183
```







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 06:21:43 ; Search time 19 Seconds  
(without alignments)  
1570.326 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910

Sequence: 1 MITALDLYHLTVAVPLVLA.....DILSTAYGPITSGFITCHS 573

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	139	4.8	302	1 Y082 METHH	O27435 methanobact
2	134	4.6	512	1 FUS BOVIN	Q28009 bos taurus
3	128.5	4.4	308	1 Y311 METJA	Q84337 methanococ
4	126.5	4.3	333	1 HME2 HUMAN	P19622 homo sapien
5	126	4.3	526	1 FUS HUMAN	P25637 homo sapien
6	124.5	4.3	623	1 PNT1 DROME	P51022 drosophila
7	123	4.2	1147	1 MYSB ACACA	P19706 acanthamoeb
8	119.5	4.1	1013	1 PMFG CHLTR	O84879 chlamydia t
9	119	4.1	718	1 MDGF RHIME	P56949 rhizobium m
10	117	4.0	320	1 MDCF RHIME	P56959 mus musculu
11	116	4.0	518	1 FUS MOUSE	P13054 drosophila
12	115	4.0	647	1 KNRL DROME	P46870 chlamydomon
13	114.5	3.9	776	1 KLP1 CHLRE	Q854h3 bifidobacte
14	112	3.8	969	1 FTSK BIFLO	Q8wx85 homo sapien
15	110.5	3.8	425	1 CC88 HUMAN	P33244 drosophila
16	110	3.8	1043	1 FTF1 DROME	Q99070 sorghum bic
17	109	3.7	168	1 GRP2 SORBI	O54839 mus musculu
18	109	3.7	688	1 BOND MOUSE	P13666 cryptotagus
19	109	3.7	908	1 SRCR RABIT	P40146 rattus norv
20	109	3.7	1248	1 CYA8 RAT	Q10707 mycobacteri
21	108.5	3.7	491	1 YK98 MYCTU	P17151 human cytom
22	108.5	3.7	684	1 EP84 HCMVA	P23813 mus musculu
23	108	3.7	323	1 HXDB MOUSE	Q928a0 homo sapien
24	108	3.7	1251	1 CYA8 HUMAN	P13709 drosophila p
25	107.5	3.7	1417	1 Y443 CHLEP	Q83983 treponema p
26	107	3.7	2038	1 FSH DROME	O89029 mesocricetu
27	106.5	3.7	506	1 GATA TREPA	O86164 chlamydia p
28	106	3.6	367	1 BET3 MESAU	Q09624 caenorhabdi
29	105.5	3.6	928	1 PM11 CHLEP	O75444 homo sapien
30	105.5	3.6	3178	1 YS85 CAEEL	Q9c0i9 homo sapien
31	105	3.6	403	1 WAF HUMAN	P07875 bacterioph
32	105	3.6	482	1 BHB3 HUMAN	
33	104.5	3.6	262	1 VG38 BET2	

Query Match

4.8%; Score 139; DB 1; Length 302;

## ALIGNMENTS

### RESULT 1

ID	YD82 METHH	STANDARD;	PRT;	302 AA.
AC	O27435;			
DT	15-JUL-1998 (Rel. 36, Created)			
DT	15-JUL-1998 (Rel. 36, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Hypothetical protein MTH1382.			
GN	MTH1382.			
OS	Methanobacterium thermoautotrophicum.			
OC	Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;			
OC	Methanobacteriaceae; Methanothermobacter.			
OX	NCBI_TaxID=187420;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Delta H.			
RX	MEDLINE=98037514; PubMed=9371463;			
RA	Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,			
RA	Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,			
RA	Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,			
RA	Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,			
RA	Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,			
RA	McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,			
RA	Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;			
RT	"Complete genome sequence of Methanobacterium thermoautotrophicum			
RT	deltaH: functional analysis and comparative genomics."			
RL	J. Bacteriol. 179:7135-7155(1997).			
CC	- - SUBCELLULAR LOCATION: Integral membrane protein (Potential).			
CC	- - SIMILARITY: TO M.JANNASCHII MJ1031.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL; AB000901; AAB85859.1; -			
DR	PIR; H69050; H69050.			
DR	InterPro; IPR004776; Auxin_eff.			
DR	Pfam; PF03547; Auxin_eff; 1.			
DR	TIGRFAms; TIGR00946;_2a69; 1.			
KW	Hypothetical protein; Transmembrane; Complete proteome.			
FT	TRANSMEM 10 30 POTENTIAL.			
FT	TRANSMEM 65 85 POTENTIAL.			
FT	TRANSMEM 102 122 POTENTIAL.			
FT	TRANSMEM 130 150 POTENTIAL.			
FT	TRANSMEM 162 182 POTENTIAL.			
FT	TRANSMEM 190 210 POTENTIAL.			
FT	TRANSMEM 224 244 POTENTIAL.			
FT	TRANSMEM 251 271 POTENTIAL.			
FT	TRANSMEM 282 302 POTENTIAL.			
SQ	SEQUENCE 302 AA; 9D3BLBA617AD42C0 CRC64;			

Query Match

4.8%; Score 139; DB 1; Length 302;

Best Local Similarity 26.0%; Pred. No. 0.013;  
Matches 52; Conservative 40; Mismatches 80; Indels 28; Gaps 8;

QY 379 GADKG---GPTLSKLSNSTAQLYPKDGG-----EGRAAAMVPPASVM--TRLLILIM--- 426  
DB 91 GLDSGKTLGTLVAAAMNNGSGFTGYPTAGIFGSELVRAIFDYDTTLTFTSLGLLSHI 150  
QY 427 -----VWRKLRNPNTYSSLIGVWMSIVSYRWGIEPPIARSISILSDAGLGMAMF 478  
DB 151 SGEGSRIKRAVFPPLMAFLGLVFNL---WG--LPTGIAGTILGYLSGAAPVLLMI 204  
QY 479 SLGLFWALOPRIACNKLAAATMGVRFVAGFAVMAASIAVGLGVLHLLTAIVQAALPQ 538  
DB 205 SLGL--TLNFRFLRHSVADATLVSLRLISPLMAAGISYVLAFLGLNFSTVLEASMP 262  
QY 539 GIVPFFVFAKEYGVHPDILST 558  
DB 263 AMLAAVLAINDLDVLDVSS 282

RESULT 2  
FUS\_BOVIN STANDARD; PRT; 512 AA.  
AC Q28009;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE RNA-binding protein FUS (Pigpen protein).  
OS Bos.  
GN Bos taurus (Bovine).  
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP TISSUE=Aorta;  
RC MEDLINE=96175600; PubMed=8631501;  
RA Alligro M.C., Alligro M.A.;  
RT "A nuclear protein regulated during the transition from active to  
RL quiescent phenotype in cultured endothelial cells."  
RL Dev. Biol. 174:288-297(1996).  
RN [2]  
RP CARBOHYDRATE-BINDING DOMAIN.  
RX MEDLINE=20160719; PubMed=10694442;  
RA Alligro M.C.;  
RT "A C-terminal carbohydrate-binding domain in the endothelial cell  
RT regulatory protein, pigpen: new function for an EWS family member."  
RL Exp. Cell Res. 255:270-277(2000).  
CC -!- FUNCTION: Binds both single-stranded and double-stranded DNA and  
CC promotes ATP-independent annealing of complementary single-  
CC stranded DNAs and D-loop formation in superhelical double-stranded  
CC DNA. May play a role in maintenance of genomic integrity (By  
CC similarity).  
CC -!- SUBUNIT: Component of nuclear riboprotein complexes. Binds SF1 (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Nuclear, exhibits diffuse staining  
CC throughout (excluding nucleoli), together with a small number of  
CC intensely stained focal points, or granules, and punctate staining  
CC along the nuclear envelope.  
CC -!- DOMAIN: The C-terminal domain binds carbohydrates.  
CC -!- SIMILARITY: Contains 1 RANBP2-type zinc finger.  
CC -!- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.  
CC -!- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

EMBL; U26024; AAC13543.1; --  
HSP; P09651; IHAL.  
DR InterPro; IPR000504; RNA rec mot.  
DR InterPro; IPR001876; Znf\_RanGDP.  
DR Pfam; PF00076; xrm; 1.  
DR Pfam; PF00641; zf-RanBP; 1.  
DR SMART; SM00360; RRM; 1.  
DR SMART; SM00347; ZNF\_RZ; 1.  
DR PROSITE; PS0102; RRM; 1.  
DR PROSITE; PS00030; RRM\_RNP; 1.  
DR PROSITE; PS01358; ZF\_RANBP2\_1; 1.  
DR PROSITE; PS01019; ZF\_RANBP2\_2; 1.  
KW RNA-binding; DNA-binding; Nuclear protein; Repeat; Zinc-finger; Zinc;  
KW Metal-binding.  
FT DOMAIN 1 164 GLN/GLY/SER/TYR-RICH.  
FT DOMAIN 165 253 GLY-RICH.  
FT DOMAIN 271 357 RNA-BINDING (RRM).  
FT DOMAIN 357 512 ARG/GLY-RICH.  
FT ZN\_FING 408 439 RANBP2-TYPE.  
SQ SEQUENCE 512 AA; 52240 MW; 3652329C044F1386 CRC64;  
Query Match 4.6%; Score 134; DB 1; Length 512;  
Best Local Similarity 23.6%; Pred. No. 0.056;  
Matches 56; Conservative 27; Mismatches 78; Indels 76; Gaps 11;  
QY 197 EAEPDGVAGAVSSRGSDAGRVRTVVRKSTSRSEAAASHSHSQTWQPRVSNLSG---- 252  
DB 69 QSAFGYGSAGYGSS-----QSSQSSYGQSSYFGYQ--QFAPSTSGSYGS 115  
QY 253 -----VEIYSLSSRNTPRGSSFNHADFFNIVGAAAKGG 288  
DB 116 SSQSSGYGQGGYGGQGGYGGQGGYGGQGGYGGQGGYGGQGGYGGQGGYGGQGGYGG 172  
QY 289 GAAG-DEEKACGGGGGGGHPQQAQVAPAKRDLHMLVWSSASPVSEAAVHVFAGG 347  
DB 173 GSYGQDQPSMSSGGGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGGYGG 217  
QY 348 ADHADVLAKGAQAYDEYGRDDYSRTKNGSGGADKGGPTLSKLSNSTALYPKDDG 404  
DB 218 YN-----RSSGGYEPGRGG-GRGGGGGGGSDRGG--FNKFGG-----PDQG 258  
RESULT 3  
YA31\_METJA STANDARD; PRT; 308 AA.  
ID YA31\_METJA  
AC Q58437;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ1031.  
GN MJ1031.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcales; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=868987;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., FitzGerald L.M., Adams M.D., Reich C.I.,  
RA Kierlavage A.R., Dougherty B.A., Tomb J.-F., Merrick J.M., Glodek A.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Weidman J.F., Nguyen D.,  
RA Scott J.L., Geoghegan N.S.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Utterback T.R., Kelley J.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Cotton M.D., Roberts K.M., Smith H.O., Woese C.R., Venter J.C.;  
RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;  
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RT jannaschii."  
RL Science 273:1058-1073(1996).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -!- SIMILARITY: TO M.THERMOAUTOTROPHICUM MTH1382.  
CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; U67545; AAB99035.1; -  
 PIR; F64428; F64428.  
 TIGR; M1031; -  
 InterPro; IPR004776; Auxin\_eff.  
 Pfam; PF03547; Auxin\_eff; 1.  
 TIGRPFAMs; TIGR00946; 2a69; 1.  
 KW Hypothetical protein; Transmembrane; Complete proteome.  
 FT TRANSMEM 5 26 POTENTIAL.  
 FT TRANSMEM 31 51 POTENTIAL.  
 FT TRANSMEM 63 83 POTENTIAL.  
 FT TRANSMEM 100 120 POTENTIAL.  
 FT TRANSMEM 128 148 POTENTIAL.  
 FT TRANSMEM 162 182 POTENTIAL.  
 FT TRANSMEM 195 215 POTENTIAL.  
 FT TRANSMEM 221 241 POTENTIAL.  
 FT TRANSMEM 257 277 POTENTIAL.  
 FT TRANSMEM 287 307 POTENTIAL.  
 SQ SEQUENCE 308 AA; 33770 MW; 02A915806E4441AE CRC64;

Query Match 4.4%; Score 128.5; DB 1; Length 308;  
 Best Local Similarity 22.5%; Pred. No. 0.072;  
 Matches 46; Conservative 46; Mismatches 87; Indels 25; Gaps 5;

QY 383 GGTLSKLSNSTAQLYPKDDG----EGRAAVAMPASVWTRLLINW-----R 429  
 DB 97 GGLIVSLMGNLTGFLGYVALGWFGEGLRAIFCDLGGVFATMLLGYVYGVIRFGKGRDK 156  
 QY 430 KLIRNPTNTSYSLGVVMSLVSRNGIEM---PAITARSISILSDAGLGMAMFSLGLFWAL 486  
 DB 157 SILKMAKPPPLTGILSIILVFFGKLYIPSPILKSLNLSATVPLINWLSGL--SL 214  
 QY 487 QPRIACGNKLAIAIANGRVFAGPAVMAAIAVGLRVGLVHLIAVQAALPQGVIPVFA 546  
 DB 215 SPKALKFGVFWGIIASIRFIVSPATATLSELINIKLEKNVLLVSSMPSAMMTLVLG 274  
 QY 547 KEYGVHPDILSTAYGPITSHGPI 570  
 DB 275 TLVELDIKLIASSI-----FIT 291

RESULT 4  
 HME2\_HUMAN STANDARD; PRT; 333 AA.  
 AC P19622;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Homeobox protein engrailed-2 (Hu-En-2).  
 GN EN2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93185339; PubMed=1363401;  
 RA Logan C., Hanks M.C., Noble-Topham S., Nallainathan D.,  
 RA Provart N.J., Joyner A.L.;  
 RT "Cloning and sequence comparison of the mouse, human, and chicken  
 RT engrailed genes reveal potential functional domains and regulatory  
 RT regions."  
 RL Dev. Genet. 13:345-358(1992).  
 RN [2]  
 RP REVISION TO 229.  
 RA Logan C., Hanks M.C., Noble-Topham S., Nallainathan D., Provart N.J.,

Joyner A.L.;  
 Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 230-333 FROM N.A.  
 RX MEDLINE=89233109; PubMed=2565873;  
 RA Poole S.J., Law M.L., Kao F.I., Lau Y.-F.C.;  
 RI "Isolation and chromosomal localization of the human En-2 gene."  
 RT Genomics 4:225-231(1989).  
 CC -|- SUBCELLULAR LOCATION: Nuclear.  
 CC -|- SIMILARITY: Belongs to the engrailed homeobox family.  
 CC -|- SIMILARITY: Contains 1 homeobox domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L12701; AAA533504.2; -  
 EMBL; L12700; AAA533504.2; JOINED.  
 EMBL; J03066; AAF68670.1; -  
 FIR; E48423; E48423.  
 HSP; P02836; 3HDD.  
 DR TRANSFAC; T02019; -  
 DR MIM; 131310; -  
 DR GO; GO:0007275; P:development; TAS.  
 DR InterPro; IPR000747; Engrailed.  
 DR InterPro; IPR001356; Homeobox.  
 DR InterPro; IPR000447; HTH\_lambrepreser.  
 DR Pfam; PF00046; homeobox; 1.  
 DR PRINTS; PR00026; ENGRAILED.  
 DR PRINTS; PR00024; HOMEBOX.  
 DR PRINTS; PR00031; HTHREPRESSR.  
 DR ProDom; PD000010; Homeobox; 1.  
 DR SMART; SM00389; HOX; 1.  
 DR PROSITE; PS00027; HOMEBOX\_1; 1.  
 DR PROSITE; PS00071; HOMEBOX\_2; 1.  
 DR PROSITE; PS00033; ENGRAILED; 1.  
 KW Homeobox; DNA-binding; Developmental protein; Nuclear protein.  
 FT DNA\_BIND 244 303 HOMEBOX.  
 SQ SEQUENCE 333 AA; 34210 MW; ACF5399B383D6257 CRC64;

Query Match 4.3%; Score 126.5; DB 1; Length 333;  
 Best Local Similarity 28.1%; Pred. No. 0.11; Indels 57; Gaps 10;  
 Matches 62; Conservative 18; Mismatches 84; Indels 57; Gaps 10;

QY 167 DQFPGAASIVSPRVDSDVVSILARGDVELEAEAPDGAVAGAVSSRG-GDAGRVRVTXK 225  
 DB 5 DPKPGENAAAV-----EGORQPESSPGGGGGGGSSPGADTGRRALMLP 51  
 QY 226 STSSRSEACSHSHSQTMQPRVSNLSGVETYSLSQSSRNPTPRGSSFNHADFFNIVGAAK 285  
 DB 52 AV---LQAPGNHQPH-----RITNFFDNLIRPEFR-----RKDAGTCCAGAGG 94  
 QY 286 GGGGGAAGDE--EKACCGGGGGHSPQPAVAVPAKXKDLHMLVWSSSASVPSRAAVHVP 343  
 DB 95 GRGGGAGGEGGASGAEGGGAGGAGGSEQ-----LLSGSGSRPRQNPCCAP-- 137  
 QY 344 GAGGADHADVLAKGAQAYDEVGRDDYSSRTKNGSGGADKGG 384  
 DB 138 GAGG-----PLPAAGS---DSPGDEGGGSKTSLHGNAKKG 171

RESULT 5  
 FUS\_HUMAN STANDARD; PRT; 526 AA.  
 ID FUS\_HUMAN  
 AC P35637;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE RNA-binding protein FUS (Oncogene FUS) (Oncogene TLS) (Translocated in  
DE liposarcoma protein) (POMP75) (75 kDa DNA-pairing protein).  
GN FUS OR TLS.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RN SEQUENCE FROM N.A. (ISOFORM LONG).  
RX MEDLINE=93288139; PubMed=8510758;  
RA Crozat A., Aman P., Mandahl N., Ron D.;  
RT "Fusion of CHOP to a novel RNA-binding protein in human myxoid  
RT liposarcoma.";  
RL Nature 363:640-644(1993).  
RN [2]  
RN SEQUENCE FROM N.A. (ISOFORM SHORT).  
RX MEDLINE=93350637; PubMed=7503811;  
RA Rabbitts T.H., Forster A., Larson R., Nathan P.;  
RT "Fusion of the dominant negative transcription regulator CHOP with a  
RT novel gene FUS by translocation t(12;16) in malignant liposarcoma.";  
RL Nat. Genet. 4:175-180(1993).  
RN [3]  
RN SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).  
RX MEDLINE=99013873; PubMed=9795213;  
RA Morohoshi F., Ootsuka Y., Azai K., Ichikawa H., Mitani S.,  
RA Munakata N., Ohki M.;  
RT "Genomic structure of the human RBP56/HTAF1168 and FUS/TLS genes.";  
RL Gene 221:191-198(1998).  
RN [4]  
RN SEQUENCE FROM N.A. (ISOFORM LONG).  
RP TISSUE=Lung, and Lymph;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loughlano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalius D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
RN SEQUENCE OF 235-244; 307-312; 335-345 AND 349-357, AND  
RN CHARACTERIZATION.  
RX MEDLINE=20036580; PubMed=10567410;  
RA Baechto H., Kuroda M., Sok J., Ron D., Lopez B.S., Akhmedov A.T.;  
RT "Human 75-kDa DNA-pairing protein is identical to the pro-oncoprotein  
RT TLS/FUS and is able to promote D-loop formation.";  
RL J. Biol. Chem. 274:34337-34342(1999).  
RN [6]  
RN SEQUENCE OF 265-276; 317-331 AND 335-357, AND IDENTIFICATION.  
RX MEDLINE=99369251; PubMed=10442642;  
RA Bertrand P., Akhmedov A.T., Delacote F., Durrbach A., Lopez B.S.;  
RT "Human POMP75 is identified as the pro-oncogene TLF/FUS: both  
RT POMP75 and POMP100 DNA homologous pairing activities are associated  
RT to cell proliferation.";  
RL Oncogene 18:4515-4521(1999).  
RN [7]  
RN CHROMOSOMAL TRANSLOCATION.  
RX MEDLINE=94243799; PubMed=8187069;  
RA Ichikawa H., Shimizu K., Hayashi Y., Ohki M.;  
RT "An RNA-binding protein gene, TLS/FUS, is fused to ERG in human

RT myeloid leukemia with t(16;21) chromosomal translocation.";  
RL Cancer Res. 54:2865-2868(1994).  
RN [8]  
RN INTERACTION WITH SFL.  
RX MEDLINE=98325009; PubMed=9660765;  
RA Zhang D., Paley A.J., Childs G.;  
RT "The transcriptional repressor ZFM1 interacts with and modulates the  
RT ability of EMS to activate transcription.";  
RL J. Biol. Chem. 273:18086-18091(1998).  
CC -1- FUNCTION: Binds both single-stranded and double-stranded DNA and  
CC promotes ATP-independent annealing of complementary single-  
CC stranded DNAs and D-loop formation in superhelical double-stranded  
CC DNA. May play a role in maintenance of genomic integrity.  
CC -1- SUBUNIT: Component of nuclear riboprotein complexes. Interacts  
CC with ILF3 and SFL.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- ALTERNATIVE PRODUCTS:  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=Long;  
CC IsoId=P35637-1; Sequence=Displayed;  
CC Name=Short;  
CC IsoId=P35637-2; Sequence=VSP\_005798;  
CC -1- TISSUE SPECIFICITY: Ubiquitous.  
CC -1- DISEASE: Involved in a form of malignant myxoid liposarcoma  
CC through a chromosomal translocation t(12;16)(q13;p11) that  
CC involves DIT3 and FUS.  
CC -1- DISEASE: Involved in a form of acute myeloid leukemia (AML)  
CC through a chromosomal translocation t(16;21)(p11;q22) that  
CC involves FUS and ERG.  
CC -1- SIMILARITY: Contains 1 RANBP2-type zinc finger.  
CC -1- SIMILARITY: Contains 1 RNA recognition motif (RRM) domain.  
CC -1- SIMILARITY: BELONGS TO THE TET FAMILY OF RNP PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
EMBL; S62140; AB27102.1; -  
EMBL; S62138; AB27103.1; ALT SEQ.  
EMBL; X71427; CAAS0558.1; ALT SEQ.  
EMBL; X71428; CAAS0559.1; ALT SEQ.  
EMBL; AF071213; AAC35285.1; -  
EMBL; AF071213; AAC35284.1; -  
EMBL; BC000402; AAH00402.1; -  
EMBL; BC002459; AAH02459.1; -  
PIR; S33799; S33799.  
HSSP; P09651; IHA1.  
DR Genew; HGNC:4010; FUS.  
DR GK; P35637; -  
DR MIM; 137070; -  
DR MIM; 151900; -  
DR GO; GO:0005634; C:nucleus; TAS.  
DR GO; GO:0003723; F:RNA binding; TAS.  
DR InterPro; IPR000504; RNA\_rec\_mot.  
DR InterPro; IPR001876; Znf\_RangDP.  
DR Pfam; PF00076; rrm; 1.  
DR Pfam; P00641; zf-RanBP; 1.  
DR SMART; SMO0360; RRM; 1.  
DR SMART; SMO0547; Znf RBZ; 1.  
DR PROSITE; PS0102; RRM; 1.  
DR PROSITE; PS00030; RRM\_RNP\_1; FALSE\_NEG.  
DR PROSITE; PS01358; ZF\_RANBP2\_1; 1.  
DR PROSITE; PS01019; ZF\_RANBP2\_2; 1.  
KW Proto-oncogene; RNA-binding; DNA-binding; Nuclear protein; Repeat;  
KW Alternative splicing; Chromosomal translocation; Zinc-finger; Zinc;  
KW Metal-binding.  
FT DOMAIN 1 165 GLN/GLY/SER/TYR-RICH.  
FT DOMAIN 166 267 GLY-RICH.  
FT DOMAIN 285 371 RNA-BINDING (RRM).

FT DOMAIN 371 526 ARG/GLY-RICH.  
 FT ZN\_FING 422 453 RANP2-TYPE  
 FT SITE 266 BREAKPOINT FOR TRANSLOCATION TO FORM  
 FT VARSPLIC 64 65 FUS/TLS-CHOP ONCOGENE.  
 FT CONFLICT 338 338 TG -> S (in isoform Short).  
 FT SEQUENCE 526 AA; 53426 MW; 88C8E263B7905549 CRC64;  
 Query March 4.3%; Score 126; DB 1; Length 526;  
 Best Local Similarity 27.1%; Pred. No. 0.21;  
 Matches 56; Conservative 18; Mismatches 75; Indels 58; Gaps 10;  
 Qy 208 AVSRGGDAGVRVTVRKSTSSREAAASHSHSOTMOPRVNLSGVBIYSLQSSRNPTPR 267  
 Db 105 APSSTSGSYG-----SSSQSSSYGQPQSGSYSQPSYGGQ--QSYGQQQSYNP-PQ 153  
 Qy 268 GSSFNHADFFNIVGAAGKGGGAAGDEKACGGGGGGHSPQPAVAVPAKRDLEMLVW 327  
 Db 154 G--YQQNQYN-----SSSGGGG-----GGGGGGNYGDDQ----- 181  
 Qy 328 SSSASPSVERAAVHFGAGGADHADVLAKGAQAYDEYGRDDYSRTKNGSGGADKGGPTL 387  
 Db 182 SSMSSG-----GSGGCGYGNQDSGGSGGGYGGQDGRGGRGSGSGGGGGGGG- 230  
 Qy 388 SKLGSNSTAQLY-PKDDGEGRAAAVAM 413  
 Db 231 --GYNRSSGCGYEPGRGGRGGRGGGM 254  
 RESULT 6  
 ID PNT1 DROME STANDARD; PRT; 623 AA.  
 AC P51022; P19420;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE ETS-like protein pointed, isoform P1 (D-ETS-2).  
 GN PNT OR ETS58B OR ETS2 OR CG17077.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 CX NCBI\_TaxID=7227;  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=20196006; PubMed=10731132;  
 RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe C., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Boutin J., Brockstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Paolis J.M., Delcher A., Deng X., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattet B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Swigras R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195 (2000).  
 RN [3]  
 RN REVISIONS, AND ALTERNATIVE SPLICING.  
 RP MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review."  
 RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).  
 RN [4]  
 RN SEQUENCE OF 456-613 FROM N.A.  
 RP MEDLINE=92249640; PubMed=1577186;  
 RX Chen T., Bunting M., Karim F.D., Thummel C.S.;  
 RA "Isolation and characterization of five Drosophila genes that encode  
 RT an ets-related DNA binding domain."  
 RL Dev. Biol. 151:176-191(1992).  
 RN [5]  
 RN SEQUENCE OF 445-603 FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL  
 RP STAGE.  
 RX MEDLINE=88196618; PubMed=2834248;  
 RA Pribyl L.J., Watson D.K., McWilliams M.J., Ascione R., Papas T.S.;  
 RT "The Drosophila ets-2 gene: molecular structure, chromosomal  
 RT localization, and developmental expression."  
 RL Dev. Biol. 127:45-53(1988).  
 CC -!- FUNCTION: Required for glial-neuronal cell interactions at the  
 CC ventral midline which are necessary for the proper elaboration of  
 CC commissures in the embryonic CNS.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=P1; Synonyms=C;  
 CC IsoId=P51022-1; Sequences=Displayed;  
 CC Name=P2; Synonyms=B;  
 CC IsoId=P51023-1; Sequences=External;  
 CC -!- TISSUE SPECIFICITY: Expressed in a complex dynamic pattern in  
 CC early embryos, including the midline and midline glial cells.  
 CC -!- DEVELOPMENTAL STAGE: Expressed throughout development with lower  
 CC levels during larval development.  
 CC -!- SIMILARITY: Belongs to the ETS family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X69166; CAA48916.1; -

```

DR EMBL; AF003742; AAN13943.1; -.
DR EMBL; M88472; AAC34200.1; -.
DR EMBL; M20408; AAA28521.1; -.
DR PIR; S33167; S33167.
DR PIR; S33168; S33168.
DR HSP; P14921; 2STT.
DR FlyBase; FBgn0003118; pnt.
DR GO; GO:0005026; F-type II transforming growth factor-beta rec. .; NAS.
DR GO; GO:0006916; P:anti-apoptosis; IGI.
DR GO; GO:0007173; P:EGF receptor signaling pathway; NAS.
DR GO; GO:0007362; P:terminal region determination; IGI.
DR GO; GO:0007179; P:FGFbeta receptor signaling pathway; NAS.
DR GO; GO:0008293; P:torso signaling pathway; IGI.
DR GO; GO:0007424; P:tracheal system development (sensu Insecta); NAS.
DR InterPro; IPR000418; Ets.
DR InterPro; IPR002341; HSF_ETS.
DR Pfam; PF00178; Ets; 1.
DR PRINTS; PR00454; ETSDOMAIN.
DR SMART; SMO0413; ETS; 1.
DR PROSITE; PS00345; ETS_DOMAIN_1; 1.
DR PROSITE; PS00346; ETS_DOMAIN_2; 1.
DR PROSITE; PS00361; ETS_DOMAIN_3; 1.
KW DNA-binding; Nuclear Protein; Developmental protein;
KW Alternative splicing.
FT DOMAIN 38 46 POLY-SER.
FT DOMAIN 47 50 POLY-ASN.
FT DOMAIN 75 88 POLY-HIS.
FT DOMAIN 103 111 POLY-GLN.
FT DOMAIN 119 124 POLY-GLN.
FT DOMAIN 128 134 POLY-GLN.
FT DOMAIN 156 159 POLY-THR.
FT DOMAIN 160 174 POLY-SER.
FT DOMAIN 246 252 POLY-GLN.
FT DOMAIN 264 267 POLY-ASN.
FT DOMAIN 294 298 POLY-ASN.
FT DOMAIN 310 320 POLY-ASN.
FT DOMAIN 323 326 POLY-ALA.
FT DOMAIN 386 389 POLY-GLY.
FT DNA_BIND 515 595 ETS-DOMAIN.
FT CONFLICT 467 467 Q -> R (IN REF. 5).
SQ SEQUENCE 623 AA; 66866 MW; P47083D960EB6480 CRC64;

```

Query Match 4.3%; Score 124.5; DB 1; Length 623;

Best Local Similarity 20.6%; Pred. No. 0.32;

Matches 80; Conservative 55; Mismatches 134; Indels 119; Gaps 15;

```

QY 161 ARALVLPQFPDGA--AAISIVSRVDSVVSLSARG-----DVELEAEPDGV 203
Db 196 AASLGIVENDMAPFVGDANAYTDSYVNFSSGYNTSNTDRINNTPPQQCQSQPTV 255
QY 204 AGAGAVSRGSDAGRVTVRKTSRSEREAACHSHSQTWQPRVSNLSVEIYSLQSRN 263
Db 256 NGSGSSNNNNNS--MLPPAVQCGNNNNNTSSNTNNSNN---NNNSGGSNNNSNAGSNN 311
QY 264 PTPRGSGFNHADFNIIVGAAA-----KGGGAAGDEEXGACGGGGGSHSPQPAVAVPA 317
Db 312 -----NNNNNNINFWAAAIIFQHLLKEPQT-----QNGNIGGYGGSGNSQNDPTDLS 361
QY 318 KRKLHMLVW--SSASPVSEARAHHVFGAGGAD---HADVLAKGAQAYDEYGRDDYSR 372
Db 362 YGLPAHLAAAYGGGSGSGTGTGRSS-----GGGGDESDYHSTISAQDHQSQSSG----- 410
QY 373 TNGSGGADKGGPTLSKLGSNSTAQLYPKDDQSEGR----- 407
Db 411 -GNSSGSGSGSGTSGNSGYLDSSEFVGSYAGNRHFDGYPEFTYDAQSFQSGMPQPT 469
QY 408 -----AAAVAMPASVMTLLI-----IMWRKLI-----RN 434
Db 470 AMDQWGAHQAHQPAATMSTLGLDKLLGLGYTTQGGVPCFTGSGPIQLMQWFLLELLDKT 529
QY 435 PNTYSSLIGVWVSL-----VSVRWGI 455
Db 530 CQSPISWTGCGWEFKLTPDDEVARRWGI 557

```

# RESULT 7

```

MYSB ACACA STANDARD; PRT; 1147 AA.
ID MYSB ACACA STANDARD; PRT; 1147 AA.
AC P19706;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Myosin heavy chain IB (Myosin heavy chain IL).
GN MIB OR MIL.
OS Acanthamoeba castellanii (Amoeba).
OC Eukaryota; Acanthamoebidae; Acanthamoeba.
CN NCBI_TaxID=5755;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90060816; PubMed=2511079;
RA Jung G., Schmidt C.J., Hammer J.A. III;
RT "Myosin I heavy-chain genes of Acanthamoeba castellanii: cloning of a
  second gene and evidence for the existence of a third isoform."
RL Gene 82:269-280(1989).
CC -!- FUNCTION: Myosin is a protein that binds to F-actin and has ATPase
  activity that is activated by F-actin.
CC -!- SUBUNIT: Myosin I heavy chain is single-headed. Dimer of a heavy
  chain and a light chain. Inability to self-assemble into filaments.
CC -!- DOMAIN: TH.1 binds directly to anionic phospholipid membranes;
  myosins I could therefore move actin relative to membranes and
  vice versa. TH.2 and SH3 bind tightly to F-actin; this together
  with the nucleotide-sensitive site in the head, allows single
  molecules of myosin I to cross-link actin filaments.
CC -!- MISCELLANEOUS: This organism expresses at least three isoforms of
  myosin I heavy-chain, encoded by genes MIA, MIB, and MIC.
CC -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -!- SIMILARITY: Contains 1 SH3 domain.

```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

EMBL; M30780; AAA27708.1; -.
DR PIR; JQ0095; MWXIB.
DR HSP; P08799; LMND.
DR InterPro; IPR000048; IQ region.
DR InterPro; IPR001609; myosin head.
DR InterPro; IPR000108; Neu_cyt_fact_2.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00612; IQ; 1.
DR Pfam; PF00683; myosin head; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00193; MYOSINHEAVY.
DR PRINTS; PR00499; P67PROX.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000355; myosin head; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00242; MYSC; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00002; SH3; 1.
DR PROSITE; ATP-binding; Phosphorylation; Multigene family; SH3 domain.
KW DOMAIN 1 670 MYOSIN HEAD-LIKE.
FT DOMAIN 671 907 TAIL HOMOLGY REGION 1 (TH.1).
FT DOMAIN 908 1089 GLY/PRO/ALA-RICH (TH.2).
FT DOMAIN 1090 1147 SH3.
FT NP_BIND 103 110 ATP (POTENTIAL).
SQ SEQUENCE 1147 AA; 124958 MW; B76DE9C076381054 CRC64;

```

Query Match

Best Local Similarity 4.2%; Score 123; DB 1; Length 1147;

Matches 60; Conservative 22; Mismatches 100; Indels 62; Gaps 9;

QY 221 VTVKTSRSEACSHSHSQTMQPRVSN-----LSGVEIYVLOSSRNPTRGSSSF 271  
 Db 861 ITYKAGNQRRLTFKSNESASQAQPSIKKSRANIQIGIATGLPKETDSSPNTWTPSGGG 920  
 QY 272 NHADFNIVGAAKGGG-GAAGDEEKGACGGGGHSPQQAQVAVPAKRDHMLVWSSS 330  
 Db 921 GYGG-----GRGGGGGRGAAGGRCGFGGGGGGYS-QPVAQAQVAVQPPVAAPSA 974  
 QY 331 -----ASPVSRAAVHVFGAGGADHADVLAQGAQYDEY 364  
 Db 975 GRGGPGMGFGAGRGFGMGFGGPGMGFGGPGMGFGGPGMGFGGPGMGFGGPGMGFG 1028  
 QY 365 GRD-----DYSSRTKNGSGGADGGPTLSK-----LGSNSTAQLYPKDGEGRAAVAM 413  
 Db 1029 GRGGPGMGFGAGRGFGMGFGGPGMGFGGPGMGFGGPGMGFGGPGMGFGGPGMG 1085  
 QY 414 PPAS 417  
 Db 1086 APAA 1089

RESULT 8  
 PMPG\_CHLTR STANDARD; PRT; 1013 AA.  
 AC 084879;  
 DT 15-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable outer membrane protein pmpG precursor (Polymorphic membrane protein G)  
 DE Protein G  
 GN PMPG OR C7871.  
 OS Chlamydia trachomatis.  
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.  
 OX NCBI\_TaxID=813;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D/UV-3/Cx;  
 RX MEDLINE=99000809; PubMed=9784136;  
 RA Stephens R.S., Kaiman S., Lemmel C.J., Fan J., Marathe R., Aravind L., Mitchell W.P., Olinger L., Tatusov R.I., Zhao Q., Koonin E.V., Davis R.W.;  
 RA "Genome sequence of an obligate intracellular pathogen of humans: Chlamydia trachomatis.";  
 RT Science 282:754-759(1998).  
 RL -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).  
 CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; AB001360; AAC68469.1; ..  
 DR F1R; G71460; G71460.  
 DR PHC1-2DPAGE; 084879; ..  
 DR InterPro; IPR006315; Autotransport.  
 DR InterPro; IPR005546; Autotransporter.  
 DR InterPro; IPR003388; Chlamydia\_PMP.  
 DR InterPro; IPR000437; Prok\_lipoprot\_S.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR Pfam; PF02415; Chlamydia\_PMP; 1.  
 DR TIGRFAMs; TIGR01414; autotrans\_bar1; 1.  
 DR TIGRFAMs; TIGR01376; POMP repeat; 6.  
 DR Outer membrane; Signal; Multigene family; Complete proteome.  
 KW SIGNAL 1 27 POTENTIAL  
 FT CHAIN 28 1013 PROBABLE OUTER MEMBRANE PROTEIN PMPG.  
 FT SEQUENCE 1013 AA; 107366 MW; F0927743C0A651DD CRC64;  
 SQ

Query Match 4.1%; Score 119.5; DB 1; Length 1013;

Best Local Similarity 21.4%; Pred. No. 1.3;  
 Matches 124; Conservative 65; Mismatches 199; Indels 191; Gaps 30;

QY 81 AVIALLL--ALASRGLSSPRALGLDWSITLFSLSITLPTNLVWG-----IPLLRGMYGAS 131  
 Db 136 SILAVLPAATWNGSQTPPTTSTPSNGTIYSKTDL-----LLNNEKFYSYMLVSGDGA 192  
 QY 132 SAGTLWQVVLQCIITWTMLFLFEYRAARALVLDOPPDGAASIV-SFRVSDVSVSLA 190  
 Db 193 DAKSLTVQGISKLCV-----PQNTAQA-----DGGACQVVTFSAMANEAPIA 236  
 QY 191 RGDVELEAPDQVAGAGAVSSRGDAGRVVTRKSTSSRSE-AACSHSHSQTMQPRVSN 249  
 Db 237 -----FIANVAGVREGGIAAVQDGGQ-----VSSSTSTEDPVVSFRNIAVEFDGNVAR 286  
 QY 250 LSGVIYSIQ-----SSRNPTRGSSFNHADFFNIVGAA-AGK 286  
 Db 287 VGG-GIYSYGNVAFLLNNGKTLFLNNVASPVYIAAEQPT-NGQASNTSDNYDGGGAIFCKN 344  
 QY 287 GGGAAAGDEBK-----ACGGGGGGHSPQQAQVAVPAKRDHMLVWSSS 330  
 Db 345 GAQAAGSNNSSVSFDGEGVVFSSNVAAGKG-----AIYAKLSV-----AN 388  
 QY 331 ASPV-----SRAAAVHVFGAG-----GADHADVLAKG-----AOAYDEYGRDDVSSR 372  
 Db 389 CGPVQFLGNIANDGGAIVLGSGLSLSADYGDIFDGNLKRKTAKENAADVNGVTVSQA 448  
 QY 373 TKNGSGG-----ADKGG-----PTLSKLSNSTAQLYPK-----DDGEGRAAAAMPPAS 417  
 Db 449 ISWGGGKITTIRAKAGHQIIFNDPIEMANGNNOQASSEPLKINDGEGTGDIVFANG- 507  
 QY 418 VMTLILIMVWRKLIIRNPNTYSSLIQVWVSVYRWGIEMPAIAR-----SISILSDAG 472  
 Db 508 -----NSTLYQNV-----IEQRIVLREKAKLSVNSLSQTG 539  
 QY 473 --LGVAMPFSLGLFMALOPRIIACGNKLAIAIANGVRFVAGPAVMAAASIAVGLRGLVLLHIA 530  
 Db 540 GSLYWEAGSTLDFVTPQ-----FQPPAANQLITLSN--LHLS 576  
 QY 531 IVQAAALPGIV-PFVFAKEYGVHPDIL-STAYGPIITSHG 567  
 Db 577 LSSLLANNAVTPNPTNPAPQDSHPAIGSTTAGSVTISG 615

RESULT 9  
 PNT2\_DROME STANDARD; PRT; 718 AA.  
 AC P51023; Q81G92; Q9VCN2;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE ETS-like protein pointed, isoform P2 (D-ETS-2).  
 GN PNT OR ETS58AB OR ETS2 OR CG17077.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.  
 RX MEDLINE=94038653; PubMed=8223245;  
 RA Klaumbt C.;  
 RT "The Drosophila gene pointed encodes two ETS-like proteins which are involved in the development of the midline glial cells.";  
 RL Development 117:163-176(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkely;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Galniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,  
 RA April J.F., Agbayan A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Buesan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Casley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Glaesermann W.,  
 RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaesermann W.,  
 RA Gladek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hades N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.B., Kodira C.B., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RA "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RN SEQUENCE OF 1-323 FROM N.A.  
 RP STRAIN=Berkeley; TISSUE=Embryo;  
 RC MEDLINE=22426066; PubMed=12537569;  
 EX Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,  
 RA George R.A., Guarini H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,  
 RA Rubin G.M., Celiker S.E.;  
 RT "A *Drosophila* full-length cDNA resource."  
 RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).  
 RN [4]  
 RN SEQUENCE OF 551-708 FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL  
 RP STAGE.  
 RC STRAIN=Canton-S; TISSUE=Larva;  
 EX MEDLINE=92249640; PubMed=1571786;  
 RA Chen T., Bunting M., Karim F.D., Thummel C.S.;  
 RT "Isolation and characterization of five *Drosophila* genes that encode  
 RL an ets-related DNA binding domain."  
 CC Dev. Biol. 151:176-191(1992).  
 CC -!- FUNCTION: Required for glial-neuronal cell interactions at the  
 CC ventral midline which are necessary for the proper elaboration of  
 CC commissures in the embryonic CNS.  
 CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event-Alternative splicing; Named isoforms=2;  
 CC Name=P2; Synonyms=B;  
 CC IsoId=P51023-1; Sequence=Displayed;  
 CC Name=P1; Synonyms=C;  
 CC IsoId=P51022-1; Sequence=External;  
 CC -!- TISSUE SPECIFICITY: Expressed in a complex dynamic pattern in  
 CC early embryos, including the midline and midline glial cells.  
 CC -!- DEVELOPMENTAL STAGE: Expressed throughout development with lower  
 CC levels during larval development.  
 CC -!- SIMILARITY: Belongs to the ETS family.  
 CC -!- SIMILARITY: Contains 1 pointed (PNT) domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X69167; CAA48917.1; --  
 DR EMBL; AR003742; AAP56125.1; --  
 DR EMBL; BT001893; AAN71882.1; ALT\_SEQ.  
 DR EMBL; M88472; AAC34200.1; --  
 DR PIR; S33168; S33168.  
 DR HSSP; P14921; 2STT.  
 DR FLYBase; FBgn00031118; pnt.  
 DR GO; GO:0005026; P-type II transforming growth factor-beta rec. .; NAS.  
 DR GO; GO:0006916; P-anti-apoptosis; IGI.  
 DR GO; GO:0007173; P-EGF receptor signaling pathway; NAS.  
 DR GO; GO:0007162; P-terminal region determination; IGI.  
 DR GO; GO:0007179; P-IGF-beta receptor signaling pathway; NAS.  
 DR GO; GO:0007424; P-tyrosine signaling pathway; IGI.  
 DR GO; GO:0007424; P-tyrosine signaling pathway; IGI.  
 DR InterPro; IPR000418; Ets.  
 DR InterPro; IPR002341; HSF ETS.  
 DR InterPro; IPR003118; SAM\_PNT.  
 DR Pfam; PF00178; Ets; 1.  
 DR Pfam; PF02198; SAM\_PNT; 1.  
 DR PRINTS; PR00454; ETSDOMAIN.  
 DR SMART; SM00413; ETS; 1.  
 DR SMART; SM00251; SAM\_PNT; 1.  
 DR PROSITE; PS00345; ETS DOMAIN 1; 1.  
 DR PROSITE; PS00346; ETS DOMAIN 2; 1.  
 DR PROSITE; PS00661; ETS DOMAIN 3; 1.  
 KW DNA-binding; Nuclear protein; Developmental protein;  
 KW Alternative splicing:  
 FT DOMAIN 166 250 POINTED.  
 FT DOMAIN 341 347 POLY-GLN.  
 FT DOMAIN 359 362 POLY-ASN.  
 FT DOMAIN 389 393 POLY-ASN.  
 FT DOMAIN 405 415 POLY-ASN.  
 FT DOMAIN 418 421 POLY-ALA.  
 FT DOMAIN 481 484 POLY-GLY.  
 FT DNA BIND 610 690 ETS-DOMAIN.  
 FT CONFLICT 133 135 DIS-> YVP (IN REF. 1).  
 SQ SEQUENCE 718 AA; 77683 MW; FDGAFD0F4BCD89C5 CRC64;  
 Query Match 4.1%; Score 119; DB 1; Length 718;  
 Best Local Similarity 20.6%; Pred. No. 0.91;  
 Matches 69; Conservative 50; Mismatches 114; Indels 102; Gaps 13;  
 Qy 197 EAPDGVAGAGVSSRGCGDAGRVTVRKSTSSRSEACSHSOTMOPRVSNLGVELY 256  
 Db 344 QSQPTVNGSGGASSNNNS-MLPFAVQSSNNNNNTSSNNNSNN---NNNSGSGNN 399  
 Qy 257 SLOSSRNPTRGSSFNHADFFNIVGAAA-----KGGGGAAGDEKGCAGGGGGHSPQP 310  
 Db 400 SNAGSNN-----NNNNNNINFMAAAATFQHHLKEEPT-----QNGNIGGYGGGSSNQ 449  
 Qy 311 QAVAVPAKRKDLHMLVW--SSSASPVSEAAVHVFGAGAD---HADVLAKGAQAYDEYG 365  
 Db 450 DPTDLSYGLPAHLAAYGGGSGGTGRSS---GGGDSRDYHSTISAQHSQQSSG 505  
 Qy 366 RDIYSRTKNGSGGADKGGPTLSKLGSNSTAQLYKPDDEGR----- 407  
 Db 506 -----GNGSGGAGSGTNGSNGYLDSSSEFYSYAGRNRFHDGYPPEFTPYDAQSFQ 557  
 Qy 408 -----AAVAMPAPASVTRLL-----IMVWRKLI 432  
 Db 558 SMCQPQTAMDQGAHAHQHPARYMSTGLDKLGLGGTTQGGVPCFTGSGFIQLWQLL 617  
 Qy 433 -----RNPNYSSLIQVWVSL-----VSYRWGI 455  
 Db 618 ELLDKTCSFISWTGDGWEFKLTDPEVARRWGI 652  
 RESULT 10  
 MDCF\_RHIME



```

Db 70 QAPQGYSTGYGSS-----QSSQSSYQQSSYGYGQ--OPAFSSYSGSYGG 116
QY 257 SIQSRNPTPRG-----SSFNHADFF---NIVGAARKGSGGAAG- 292
Db 117 SSQSSSYGQPGQSGYGGQSSYGGQSSYNNPPQGYGQQNQYNSSSGGGGGGGG 176
QY 293 ----DEEKACGGGGGSHSPQCAVAVPAKRDLEHLVWSSASPVSRAAHVFGAGGA 348
Db 177 NYGQDQSSMSGGGGGGGYNQDQS-----GGGGG 205
QY 349 DHA-----DVLAKGAQAYDEYGRDD--YSRTKNGS-----GGADKGQPTLSKLGNS 395
Db 206 GYGGGQDGRGGRGCGGGGYNRSSGYEPRGCGGCGRGCGGWSDRGG--FNKFGG--- 260
QY 396 AGLYKDDG 404
Db 261 ----PRDQG 265

RESULT 12
KNRL_DROME
ID KNRL_DROME STANDARD; PRT; 647 AA.
AC F13054; Q9VPC8;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Knirps-related protein.
GN KNRL OR NR0A2 OR CG4761.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89057149; PubMed=2848202;
RA Oro A.E., Ong E.S., Margolis J.S., Posakony J.W., McKewon M.,
RA Evans R.M.;
RT "The Drosophila gene knirps-related is a member of the
RL steroid-receptor gene superfamily.",
RL Nature 336:493-496(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Anantides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laslo P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattet B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleeb J.M.,
RA Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,

```

```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamis I., Simpson M., Skubecki M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [3]
RP REVISIONS.
RX MEDLINE=22426069; PubMed=12537572;
RA Miera S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.O.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Kronmiller B., Pacleeb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the nuclear hormone receptor family. NRO
CC subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X14153; CAA32365.1; --
CC EMBL; AE003591; AAP51627.2; --
CC EMBL; AY075383; AAL68221.1; --
CC PIR; S06450; S06450.
CC HSSP; P20393; 1A6Y.
CC TRANSFAC; T02773; --
CC FlyBase; FBgn0001323; knr1.
CC GO; GO:0004879; F:ligand-dependent nuclear receptor activity; NAS.
CC GO; GO:0007088; P:regulation of mitosis; IMP.
CC InterPro; IPR001628; Znf_C4steroid.
CC Pfam; PF00105; zf-C4; 1.
CC PRINTS; PR00047; STROIDFINGER.
CC ProDom; PD000035; Znf_C4steroid; 1.
CC SMART; SM00399; Znf_C4; 1.
CC PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
CC Receptor; Transcription regulation; DNA-binding; Nuclear protein;
CC Zinc-finger; Developmental protein.
CC DINA BIND 14 80 NUCLEAR RECEPTOR-TYPE.
CC ZN FING 14 34 C4-TYPE.
CC ZN FING 51 75 C4-TYPE.
CC DOMAIN 121 262 GLY-RICH.
CC FT DOMAIN 577 597 ASN-RICH.
CC SEQUENCE 647 AA; 68322 MW; E466FA081DAACDC8 CRC64;
Query Match 4.0%; Score 115; DB 1; Length 647;
Best Local Similarity 21.7%; Pred. No. 1.5;
Matches 68; Conservative 33; Mismatches 99; Indels 114; Gaps 15;

```

QY 200 PDGVAGAGAYSSRGDAGRVRVTVRKSTSSRSEAAAC-----SHSHSQTMOPRV 247  
 Db 160 PGYVGLYAVANAG-----SSRSKEELMWGLDGSVEYSGHHPVVASPSV 206  
 QY 248 SNL-----SGVIYSLQSSRNTPRGSSFNADFFNIVGAAAKGGGAGDE--EKGAC 299  
 Db 207 SSPDSHNSDSVSVSV-----RGNPLH-----LGGKNSGGSSGADGSHSGG 252  
 QY 300 GGGGGGHSPP-OPQAVAVPAKRKDL-----HML 325  
 Db 253 GGGGGGVTPGRP-----PQWKDLSPLPLPPGLASMFVMPAPLPPSHLLPGYHPA 307  
 QY 326 VWS-----SASPVSRAAV-----HVFAGAGADHADVLAQAAY-----DEVGRDDYS 370  
 Db 308 LYSHQGLLPTPEQQAQAAVAAVQELFNSSGAGOR--FAPGTSPFANHQHKEEDQP 365  
 QY 371 SRKNSGGGADK-----GGFTLSKLSNSTAQLPKDQDEGRAAAVAMPAPVMTLILI 425  
 Db 366 APARSPTTHANNHLLTNGGADELTKRFYLDVAVLSQQSQSPPTTKLPPhSKQDYSISA 425  
 QY 426 MVWRKLIRNPNTYS 439  
 Db 426 LV-----TPNSES 433

RESULT 13  
 ID\_KLPI\_CHLRE STANDARD; PRT; 776 AA.  
 AC P46870;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Kinesin-like protein KLPI.  
 GN KLPI.  
 OS Chlamydomonas reinhardtii.  
 OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;  
 OC Chlamydomonadaceae; Chlamydomonas.  
 OX NCBI\_TaxID=3055;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=21gr;  
 RX MEDLINE=94266963; PubMed=8207060;  
 RA Bernstein M., Beech P.L., Katz S.G., Rosenbaum J.L.;  
 RT "A new kinesin-like protein (Klp1) localized to a single microtubule  
 of the Chlamydomonas flagellum.";  
 RL J. Cell Biol. 125:1313-1326(1994).  
 CC -!- FUNCTION: May play a role in rotation or twisting of the central  
 pair microtubules of the flagella axoneme.  
 CC -!- SUBCELLULAR LOCATION: BOUND TO THE CENTRAL PAIR MICROTUBULE C2 OR  
 TO PROJECTIONS THAT ARE ATTACHED TO C2.  
 CC -!- SIMILARITY: Belongs to the kinesin-like protein family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; X78589; CAA5326.1; -;  
 DR PIR; A53953; A53953.  
 DR HSPSP; P17119; 3KAR.  
 DR InterPro; IPR001752; kinesin\_motor.  
 DR Pfam; PF00225; kinesin; 1.  
 DR PRINTS; PR00380; KINESINHEAVY.  
 DR SMART; SM00129; KISC; 1.  
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.  
 DR PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.  
 DR PROSITE; PS0067; KINESIN MOTOR DOMAIN3; 1.  
 KW Motor protein; Microtubule; ATP-binding; Coiled coil.  
 FT DOMAIN 1 342 KINESIN MOTOR (BY SIMILARITY).  
 FT DOMAIN 348 371 COILED COIL (POTENTIAL).  
 FT DOMAIN 579 657 COILED COIL (POTENTIAL).  
 FT DOMAIN

FT DOMAIN 658 776 GLOBULAR (POTENTIAL).  
 FT NP\_BIND 91 98 ATP (POTENTIAL).  
 SQ SEQUENCE 776 AA; 83020 MW; 48646509DE8AA71D CRC64;  
 Query'Match 3.9%; Score 114.5; DB 1; Length 776;  
 Best Local Similarity 25.2%; Pred. NO. 2;  
 Matches 65; Conservative 26; Mismatches 94; Indels 73; Gaps 12;  
 QY 195 ELEAE---PDGVAGAGAYS---SRGGDAGRVRVTVRKSTSSRSEAAACSHSHS-OTMOPRV 247  
 Db 360 ELKAEALMDTLISGKRVSYDDLTDELRELHATCRRLHGEAPEDLPADSMKRVRETF 419  
 QY 248 SNLSGVEI-----YSLQSSRNTPRGSSFNADFFNIVGAAAKGGGGA-----AGDEEK 296  
 Db 420 KALRAVHVAIKADMATOMATLRRATEEGS-----GAAARGGDSAGSGVGVVDL 468  
 QY 297 GACGGGGGHSPPQAVAVPAKRKDLHMLVWSSSASPVSRAAAVHVFG-----AGADHAD 352  
 Db 469 RATGGFTVGHAFLD---ARPPVRSSEL-----GSPGAGASGAELGEPSPGGLHAQ 517  
 QY 353 VLAKGAQAYDEYGRDDYSRTKNGSGGADKGGPTLSKLS-----NSTAQLYP 400  
 Db 518 A-----SSHTDAGSNWGD-AGPLSSPGGTRLAGIFGVSGDRNAVFRYK 560  
 QY 401 XDDGGRRAAAVAMPASV 418  
 Db 561 VDVGEGRELAASLKAASI 578

RESULT 14  
 ID\_FTSK\_BIFLO STANDARD; PRT; 969 AA.  
 AC Q8G4H3;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE DNA translocase ftsK.  
 GN FTSK OR BLI411.  
 OS Bifidobacterium longum.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;  
 OC Bifidobacteriaceae; Bifidobacterium.  
 OX NCBI\_TaxID=216816;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCC 2705;  
 RX MEDLINE=22294977; PubMed=12381787;  
 RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,  
 RA Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,  
 RA Pridmore R.D., Arigoni F.;  
 RT "The genome sequence of Bifidobacterium longum reflects its adaptation  
 to the human gastrointestinal tract.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).  
 CC -!- FUNCTION: DNA motor protein, which is both required to move DNA  
 out of the region of the septum during cell division and for the  
 septum formation. Tracks DNA in an ATP-dependent manner by  
 generating positive supercoils in front of it and negative  
 supercoils behind it (By similarity).  
 CC -!- SUBUNIT: Homohexamer. This suggests the formation of a ring  
 between the two cells at the septum that surrounds DNA (By  
 similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Located at the  
 septum. The large C-terminal part of the protein is cytoplasmic  
 (potential).  
 CC -!- SIMILARITY: Contains 1 FtsK domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC

```

DR EMBL; AE014771; AAN25210.1; -.
DR HAVAP; MF_01809; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR002543; FtsK_SpoIIIE.
DR InterPro; IPR008253; Marv.
DR Pfam; PF01580; FtsK_SpoIIIE; 1.
DR Pfam; PF01284; MARVEL; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS0901; FTSK; 1.
KW Chromosome partition; Cell division; ATP-binding; DNA-binding;
KW Transmembrane; Complete proteome.
FT TRANSMEM 73 92 POTENTIAL.
FT TRANSMEM 102 124 POTENTIAL.
FT TRANSMEM 136 158 POTENTIAL.
FT TRANSMEM 168 190 POTENTIAL.
FT TRANSMEM 192 214 POTENTIAL.
FT DOMAIN 598 798 FTSK.
FT NP_BIND 615 622 ATP (POTENTIAL).
SQ SEQUENCE 969 AA; 102650 MW; 50711A08044CB794 CRC64;

Query Match 3.8%; Score 112; DB 1; Length 969;
Best Local Similarity 20.9%; Pred. No. 4;
Matches 94; Conservative 65; Mismatches 176; Indels 114; Gaps 19;

QY 41 GINFEVALFAVPLSLFHSFISTNDPPFAMNLFPLAADTLQKVALALLAS----- 90
DB 72 GLCELLLVAVLFCASEFVSGPFGQLLHAIAGVGLMSVLPVLLAAVAFRLMRNSG 131

QY 91 RGLSPRALG-----LDWSI-----TLFSLSTLPNT-----LVNGIPLLRGMVG 129
DB 132 KGSNNPRVVTGWLLMWSICSIIDVAIAADHTGDTILQSAGLFGFLGSLP---AWG 188

QY 130 ASSAGTLMQVQVLCIITWLMFLFEY-----RAARALV-----LDQFPDGA 174
DB 189 LSNVFAIIFVW-----GLFSLMITGTHVTDLPEDARKIAKIQKPYVPMGQETDGA 244

QY 175 ASIVS-FRVSDDVSVLSLARGVELEAPDGVAGAGAVSSRGDA-----GRVVRVTR 224
DB 245 SQPNEVRVGTDTLAFADGVPSHDGDD-----GSDNDQAGDARPSLFAFLGRKSKTD 299

QY 225 KSTSRSEA-----ACSHSHSQTMQPRVNSLGSVEIYLSQSRPTFRGSGFNAADFN 278
DB 300 DKTLDKYAADPPFRAASQHGTAETFPVDPMTG-EIIGAR-----TIASSYDGRPHLS 353

QY 279 IVGAAAKGGGGAAG-----DEEKAGCGGGGGHSPQQAQAVAPAKRKDLHMLVWSSAS 332
DB 354 SPAPAADADDGASRTVITSGQTVAMPGGAVDDP-----WAPSA 395

QY 333 PVSERAAHVFGAGGADHADVLAQQAQVDEYGRDDYSSTKNGSGGADKGGPTLSKLS 392
DB 396 QAGTVLAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 454

QY 393 NSTAQLYFKDDGEGRAAAVAMPASVMT 421
DB 455 DT-----DDANR--PYQLPDLNLT 474

RESULT 15
CCG8 HUMAN
ID CCG8 HUMAN STANDARD; PRT; 425 AA.
AC QWXS5; Q9BX70; Q9BY23;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE Voltage-dependent calcium channel gamma-8 subunit (Neuronal voltage-
DE gated calcium channel gamma-8 subunit).
GN CACNG8 OR CACNG6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=21601102; PubMed=11738816;
RA Chu P.-J., Robertson H.M., Best P.M.;
RT "Calcium channel gamma subunits provide insights into the evolution of
RN this gene family.";
RL Gene 280:37-48(2001).
RP [2]
RP SEQUENCE OF 12-426 FROM N.A.
RA MEDLINE=21100909; PubMed=11170751;
RA Burgess D.L., Gefrides L.A., Foreman P.J., Noebels J.L.;
RT "A cluster of three novel Ca(2+) channel gamma subunit genes on
RN chromosome 19q13.4: evolution and expression profile of the gamma
RL subunit gene family.";
RP Genomics 71:339-350(2001).
RP [3]
RP SEQUENCE OF 1-204 FROM N.A.
RA TISSUE=Cerebellum;
RA Black J.L. III, Kryzer T.J., Lennon V.A.;
RT "Proposed Homo sapiens voltage-gated calcium channel gamma-6
RN subunit.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Thought to stabilize the calcium channel in an
CC inactivated (closed) state (By similarity).
CC -1- SUBUNIT: The L-type calcium channel is composed of five subunits:
CC alpha-1, alpha-2/delta, beta and gamma.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the PMP-22 / EMP / MP20 family. CACNG
CC subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF361354; AAL50049.1; -.
DR EMBL; AF288388; AAK20031.1; -.
DR EMBL; AF234892; AAK15019.1; ALT_INIT.
DR Genew; HGNC:13628; CACNG8.
DR MIM; 606900; -.
DR GO; GO:0005891; C: voltage-gated calcium channel complex; NAS.
DR GO; GO:0005245; F: voltage-gated calcium channel activity; NAS.
DR GO; GO:0006816; P: calcium ion transport; NAS.
DR InterPro; IPR004031; PMP22_Claudin.
DR InterPro; IPR008368; VDCCGamma.
DR InterPro; IPR008372; VDCCGammaa.
DR Pfam; PF00822; PMP22_Claudin; 1.
DR PRINTS; PR01792; VDCCGamma.
DR PRINTS; PR01796; VDCCGammaa.
DR Ionic channel; Transmembrane; Ion transport; Voltage-gated channel;
KW Calcium channel.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 129 149 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 208 228 POTENTIAL.
FT TRANSMEM 318 338 POTENTIAL.
FT DOMAIN 247 380 GLY-RICH.
FT CONFLICT 1 2 ME -> QV (IN REF. 3).
FT CONFLICT 16 16 E -> K (IN REF. 3).
FT CONFLICT 67 70 HRGG -> PPAGA (IN REF. 1).
FT CONFLICT 362 362 S -> A (IN REF. 2).
FT CONFLICT 386 386 R -> G (IN REF. 2).
FT CONFLICT 394 395 RH -> AP (IN REF. 2).
FT CONFLICT 409 409 G -> E (IN REF. 2).
SQ SEQUENCE 425 AA; 43481 MW; E1105BDAF619D3D CRC64;

Query Match 3.8%; Score 110.5; DB 1; Length 425;
Best Local Similarity 23.5%; Pred. No. 1.9;
Matches 57; Conservative 18; Mismatches 71; Indels 97; Gaps 10;

QY 247 VSNLSGVETYSLSRSNPTPR-----GSSFNEADFFNIVGAAA----- 284

```

Db	171	LSNIIGVIVISANAGEPCPKDEEKXKHYSYGMFYFGLSFLAE---VIGVLAVNIY	227
Qy	285	-----KGGGAAGDEEKAGCGGGGHS-----	307
Db	228	IERSREAHCCSRDILLKAGGAG-----GSGGSGPSAILRLPSYFRYRRSRSSSR	279
Qy	308	----PQPAVAVPA-----KRDHLMLVWSSSASFSERAAVHVFGAGGADHADVLAKG	357
Db	280	SSEPSRDSAPGGPGPGFASDISM--YTLRDPKGSVAAGLAGAGGGGGGAVGAFG	337
Qy	358	AOAYDEYGRDDYSSRTKNGSGGA---DKGGPTLSKLGSNSTAOLYPKDDGEGRAAAVAM	413
Db	338	GAAGGAGGG-----GSGGGAERDRGGSS-----GFLTLHNAFPKEAGGGVTVTTR	386
Qy	414	PPA	416
Db	387	PPA	389

Search completed: March 3, 2004, 09:07:26  
 Job time : 24 secs



```

Db 118 IPLKGMYG-NSSGSLMVQIVLQCIITWYLMFLFYRGARLIGEQPFD-TAGSIIISF 175
QY 181 RVDSDVSL-ARGDVELEAEPDGVAGAGVSRGGDAGRVVTVRKSTSSRSEAAASHSH 239
Db 176 RVDSDIISLDGREPLOTAEV-----GEDGKLHVTVRKSTSSRSEVFSHMSH 222
QY 240 ---SQTQPRVSNLGVETIYSLQSSRNPTPRGSSFNHADFFNIVGAAAKGGGGAAG-- 292
Db 223 GLNSGLSITPRPSNLTWAEIYSLQSSRNPTPRASSFNHTDFYSMV-----NKGNASPRHS 277
QY 293 -----DEKGCAGCGGGGHSPOQAVAPD-----KRDHLMLIV 326
Db 278 NFTNLQDESG--GLGVGNVPRANGSAYPAPNAGIFSPGCKKXANGKGLHMFV 335
QY 327 WSSSASPVSEAAVHVFGAG-----ADHADVLAKGAQAVDEYGRDDYSRTKNGSG 378
Db 336 WSSSASPVSE--GGLHVFKGDDYNDLGGVAHKK-----YDEGRDEFSGNRPGN 386
QY 379 GADKGGTTLKSLGNSSTAQLYKDDGEGRAAAVAMPASVMTLILIMVWKLIRNPNTY 438
Db 387 GVDKGGVPLKSGSSSTAEHLHPKSAANGAEAKPTWMPFTSVMTLILIMVWKLIRNPNTY 446
QY 439 SSLIGTVVSLVSVRWGIEMPAIARSISILSDAGLGMAMFSLGLFMALQPRIAACNKLA 498
Db 447 SSLIGTWSLVSFKNMTPALLANSIALISNAGLGMAMFSLGLFMALQPRIAACNSIA 506
QY 499 AIAKGVFVAGPAVMAAASAVGLRGVLLHIAIVQAAIPQGVFFVFAKEYGVHDPILST 558
Db 507 SFANGVRPLTGPVMAAASFAVGLRGDLRLRIAIVQAAIPQGVFFVFAKEYGVHDPILST 566
QY 559 A--YG-----PIT 564
Db 567 GVIFGMLIALPIT 579

```

## RESULT 2

```

ID O81215 PRELIMINARY; PRT; 595 AA.
AC O81215;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Auxin transport protein REH1.
GN REH1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98344010; PubMed=9679062;
RA Lueschning C., Gaxiola R.A., Grissafi P., Fink G.R.;
RT "ERL1, a root-specific protein involved in auxin transport, is
RT required for gravitropism in arabidopsis thaliana.";
RL Genes Dev. 12:2175-2187(1998).
DR EMBL; AF056027; AAC39514.1; -.
DR PIR; T02876; T02876.
DR Gramene; O81215; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; 1.
DR TIGRFAMs; TIGR00946; za69; 1.
SQ SEQUENCE 595 AA; 64717 MW; 81703D7382CEC2A8 CRC64;

```

```

Query Match 57.2%; Score 1665; DB 10; Length 595;
Best Local Similarity 58.98; Pred No. 3 2e-106;
Matches 365; Conservative 54; Mismatches 111; Indels 90; Gaps 17;
QY 1 MITADLYHYLTAVVPLVYVMTLAYGSRVWRWRIPTDQSGINRFVAFVPLLSFHTIS 60
Db 1 MITAADFYHVTAMVPLVYVMTLAYGSKVWRWRIPTDQSGINRFVAFVPLLSFHTIS 60

```

```

QY 61 TNDPFFAMNRFPLAADTLQKVAIVALLA-----LASRGLSSPRALGLDWSITLFSITLPTNT 116
Db 61 TNNPYTMNRFPLAADTLQKVAIVALLA-----SLEWTTITLFSITLPTNT 113
QY 117 LVNGIPELLRGMTGASSAGTLMVQVVLQCIITWYLMFLFYRGARLIGEQPFDGGAAS 176
Db 114 LVNGIPELLRGMTGASSAGTLMVQVVLQCIITWYLMFLFYRGARLIGEQPFD-TAAN 171
QY 177 IYVFRVDSVSL--ARGDVELEAEPDGVAGAGVSRGGDAGRVVTVRKSTSSRSEAA 234
Db 172 IASIVVDPDQVSLDGRDAIETEVK-----EDGRHVTVRSSNASRSDIY 218
QY 235 CSHSHS-QTQPRVSNLGVETIYSLQSSRNPTPRGSSFNHADFFNIVGAAAKGGGGAAGD 293
Db 219 SRRSMGFSTTTPRPSNLTWAEIYSLQSSRNPTPRGSSFNHTDFYSMVGRSSNFGAADA 278
QY 294 BEKGAC-----GGGGGH--SPQOAVAVP-----AKR 319
Db 279 VRTGATPRPSNTEDDASKPKYPLPASNAAPMAGHYPAFNPVAVSSAPKGAKAATNGQAKG 338
QY 320 KDLHMLVWSSSSSPVSEAAVHVFGAGGADHADVLA-----KGAQAYDEY-GRDDYSS 371
Db 339 EDLHMFVWSSSSSPVSD-----VFGGAPDYNDAAAVKSPKMDGAKDREDYVERDDFSF 393
QY 372 RTKNGSGGADKGGPTLSKLGNSSTAQLYKDDGEGRAAAVAMPASVMTLILIMVWKL 431
Db 394 GNR---GVMDRD-----AEAGDEKAAAGADPSKAMAAPTWMPFTSVMTLILIMVWKL 446
QY 432 IRNPNTYSSLIGVWVSLVSVRWGIEMPAIARSISILSDAGLGMAMFSLGLFMALQPRII 491
Db 447 IRNPNTYSSLIGVWVSLVSVRWGIEMPAIARSISILSDAGLGMAMFSLGLFMALQPHII 506
QY 492 ACNKLAATAMGVRFVAGPAVMAAASAVGLRGVLLHIAIVQAAIPQGVFFVFAKEYGV 551
Db 507 ACNKRVATYAMVRFVAGPAVMAAASAVGLRGVLLHIAIVQAAIPQGVFFVFAKEYGV 566
QY 552 HPDILSTA--YG-----PIT 564
Db 567 HPSILSTAVIFGMLIALPIT 586

```

## RESULT 3

```

QY 09C6B8 PRELIMINARY; PRT; 622 AA.
ID 09C6B8
AC 09C6B8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Auxin transporter splice variant b, putative (At1g73590/P6D5_2).
GN F6D5.2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=cv. Columbia;
RC MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Buehler E., Chan A., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Chung M.K., Conn L., Conway A.B., Feng J.-D., Fong B., Fujii C.Y.,
RA Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Hanes N.F., Hughes B., Huizar L.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Khan S., Khaykin E.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Kwan A., Lam B.,
RA Kim C.J., Koo H.L., Krenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Millescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,

```

RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,  
 RA Uterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,  
 RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;  
 RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis  
 RT thaliana";  
 RL Nature 408:816-820(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,  
 RA Bower L., Carninci P., Chung M.K., Goldsmith A.D., Hayaishizaki Y.,  
 RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,  
 RA Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M., Nguyen M.,  
 RA Palm C.J., Pham P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Toriumi M., Yamada K., Yu G., Shinozaki K.,  
 RA Davis R.W., Theologis A., Ecker J.R.;  
 RT "Arabidopsis cDNA clones";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Cheuk R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banh J.,  
 RA Bower L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayaishizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RT "Arabidopsis ORF clones";  
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC079676; AAG51807.1; -;  
 DR EMBL; AF372950; AAK50090.1; -;  
 DR EMBL; AY093960; AAM16221.1; -;  
 DR EMBL; G96762; G96762;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR InterPro; IPR004776; Auxin\_eff.  
 DR Pfam; PF03547; Auxin\_eff; 1.  
 DR TIGRFAMs; TIGR00946; 2a69; 1.  
 SQ SEQUENCE 622 AA; 67018 MW; 333E90DB12F62C83 CRC64;

Query Match 56.6%; Score 1647.5; DB 10; Length 622;  
 Best Local Similarity 56.2%; Pred. No. 5.4e-105;  
 Matches 366; Conservative 56; Mismatches 104; Indels 125; Gaps 18;

QY 1 MITADLYHVLTAIVPLVYVAMTAYGSRVWRKIPTDQCNGINRFVALFAVPLLSFHFI 60  
 DB 1 MITADLYHVLTAIVPLVYVAMTAYGSRVWRKIPTDQCNGINRFVALFAVPLLSFHFI 60

QY 61 TNDPFAKMLRFLAADTLQKVAVALALLASRGLSSPRALGLDWSITLPSLTLPNT 120  
 DB 61 ANNPYAMNLRFLAADTLQKVAVALALLASRGLSSPRALGLDWSITLPSLTLPNT 117

QY 121 IPLLGMYGASAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQFPDGAASIVSF 180  
 DB 118 IPLLGMYG-NFSGDLVQIVVLQCIWYTLMLFLFEYRAARALVLDQFPDGAASIVSF 175

QY 181 RYDSDVVSI-ARGDVELEAPDGVAGAVSSRGDAGRVVTRKSTSRSEACSHSH 239  
 DB 176 HVDSIDMSLDGQVLETAIEIK-----EDGKLHVTRKSNASRDIF 222

QY 240 SQTMQPRVNLGVEIYSIQSSRNPTPRGSSFNHADFNIY----- 280  
 DB 223 GLSATPRPNLTAETIYSIQSSRNPTPRGSSFNHADFNIY----- 282

QY 281 -----GAAAK-GGGGAG-----DEKGACGGGGGSHPOP----- 310  
 DB 283 KGTTPRPNYEDGGPAKTAAGTAAGRFRHYQSGGGGGGGHAYPAPNCPMSPNTGG 342

QY 311 -----QAVAPAKR-----KDLMLVWSSASPVSERAAVHVFAGGADHADVLA 357  
 DB 343 GGGTAAGKAPVVGKQDNGERDLHMFVWSSASPVS-----VFGGGGNH----- 394

QY 358 AQAYDEYGRD-----DYSSRTKNGSGGADKGGPTLSKLK-----SNSTAQLYP 400  
 DB 358 AQAYDEYGRD-----DYSSRTKNGSGGADKGGPTLSKLK-----SNSTAQLYP 400

Db 395 STATNDHQDKVSKISVPOGNSNDNQVREBEFSGNKDDSKVLATDGGNNISNKTTO--- 451  
 QY 401 KDDGEGRAAAVAMPASPATVRLILIMVWRKLRNPNTYSSLLIGVWVSLVSRWGIEMPAI 460  
 Db 452 -----AKVMPPTSVNTRILLIMVWRKLRNPNTYSSLLIGVWVSLVSRWGIEMPAI 502  
 QY 461 IARSISILSDAGLGMASFSLGFALQPRIIACGNKLAATAMGVRFVAGFAVMAAASIAV 520  
 Db 503 IAKSISILSDAGLGMASFSLGFALQPRIIACGNRRRAFAAAMRFVVGFAVMLVASYAV 562  
 QY 521 GLRGVLLHIAIQAALPQGIPIVFFFAKEYGVHPDILSTA--YG-----PIT 564  
 Db 563 GLRGVLLHVAIIQAALPQGIPIVFFFAKEYNVHPDILSTAVIFGMLIALPIT 613

RESULT 4  
 Q8H0E0 PRELIMINARY; PRT; 617 AA.  
 ID AC Q8H0E0;  
 DT 01-MAR-2003 (TREMELrel. 23, Created)  
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE PIN1-like auxin transport protein.  
 GN CS-PIN1.  
 OS Cucumis sativus (Cucumber).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Cucurbitales; Cucurbitaceae; Cucumis.  
 OC NCBI TaxID=3659;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kamada M., Yamasaki S., Fujii N., Higashitani A., Takahashi H.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Motooshi K., Seiji Y., Nobuharu F., Atsushi H., Hideyuki H.;  
 RT "Gravity-induced modification of auxin transport and distribution for  
 RT peg formation in cucumber seedlings: Possible roles for CS-AUX1 and  
 RT CS-PIN1";  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB085897; BAC41319.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR InterPro; IPR004776; Auxin\_eff.  
 DR Pfam; PF03547; Auxin\_eff; 1.  
 DR TIGRFAMs; TIGR00946; 2a69; 1.  
 SQ SEQUENCE 617 AA; 67344 MW; D61D6E421F231382 CRC64;

Query Match 56.6%; Score 1647; DB 10; Length 617;  
 Best Local Similarity 56.6%; Pred. No. 5.8e-105;  
 Matches 368; Conservative 54; Mismatches 100; Indels 128; Gaps 21;

QY 1 MITADLYHVLTAIVPLVYVAMTAYGSRVWRKIPTDQCNGINRFVALFAVPLLSFHFI 60  
 DB 1 MITADLYHVLTAIVPLVYVAMTAYGSRVWRKIPTDQCNGINRFVALFAVPLLSFHFI 60

QY 61 TNDPFAKMLRFLAADTLQKVAVALALLASRGLSSPRALGLDWSITLPSLTLPNT 116  
 DB 61 TNPFTMLRFLAADTLQKVAVALALLASRGLSSPRALGLDWSITLPSLTLPNT 113

QY 117 LVMGIPLLRGMYGASAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQFPDGAAS 176  
 DB 114 LVMGIPLLRGMYGASAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQFPDGAAS 171

QY 177 IVSFRVSDVVSIL-ARGDVELEAPDGVAGAVSSRGDAGRVVTRKSTSRSEAAC 235  
 DB 172 IVSFRVSDVVSIL-ARGDVELEAPDGVAGAVSSRGDAGRVVTRKSTSRSEAAC 218

QY 236 SHSHS-QTMQPRVNLGVEIYSIQSSRNPTPRGSSFNHADFNIYVGAAGGGGAGD- 293  
 DB 219 RRSVGLSTTPRPNLTAETIYSIQSSRNPTPRGSSFNHADFNIYVGAAGGGGAGD- 278

QY 294 -----REKGACGGGG-----CHSP----- 308  
 DB 294 -----REKGACGGGG-----CHSP----- 308

```

Db 279 YGLSASRGPTPRPSNYEEEG--CGGGKPRPHYNATTGGNANANANANVHYPAHPGMF 336
QY 309 -----QPAVAVPAKPK-----DLHMLWSSSPVSEAAVHVFGAGDADVD- 353
Db 337 SPTGSKNAOPNNAKPAKTEDGGDRLHMFVWSSASPVSDVFNHFGAHN-DQKQVR 395
QY 354 --LAQAQA--YDEYG-RDDYS-----SRTKNGSGGADKGPTLSKLSNSTAQLYPK 401
Db 396 LAVSPGKEGRRENQEBYAEERDFSGNREMMNSNNGGVGVG-----GTEKVGDIKPK 448
QY 402 DDEGGRAAVAVPPASVMTLILIMVWRKLIENPNYSSLIQVWLSVSYRWGTEMPAI 461
Db 449 -----TMEPTSVMTLILIMVWRKLIENPNYSSLIQVWLSVSYRWGTEMPAI 498
QY 462 ARSISILSDAGLGMAFMFSLGLFMALQPRIIACGNKLAIAIAMGVRFVAGPAAVMAASIAVG 521
Db 499 AKSISILSDAGLGMAFMFSLGLFMALQPRIIACGNKLAIAIAMGVRFVAGPAAVMAASIAVG 558
QY 522 LRGVLLHIAIQAALPQGIQVFPVFAKEYGVHDPDILSTA--YG-----PIT 564
Db 559 LRGVLLHIAIQAALPQGIQVFPVFAKEYGVHDPDILSTGVIFQMLVALPIT 608

RESULT 5
Q9ZSY6 PRELIMINARY; PRT; 622 AA.
AC Q9ZSY6;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative auxin efflux carrier protein.
GN PIN1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=99074388; PubMed=9856939;
RA Gailweller L., Guan C., Mueller A., Wisman E., Mendgen K.,
RA Yephremov A., Palme K.;
RT "Regulation of polar auxin transport by AtPIN1 in Arabidopsis vascular
RT tissue."
RL Science 282:2226-2230 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Gailweller L., Changhui G., Mueller A., Wisman E., Palme K.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF089084; AAD04376.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; I.
DR TIGRFAMs; TIGR00946; 2a69; 1.
SQ SEQUENCE 622 AA; 67031 MW; 8A50F9C2793E8CC5 CRC64;

```

```

Query Match 56.4%; Score 1641.5; DB 10; Length 622;
Best Local Similarity 56.1%; Pred. No. 1.4e-104;
Matches 365; Conservative 56; Mismatches 105; Indels 125; Gaps 18;

QY 1 MITALDLYHVTAVVPLVYAMTAYGSRVWRRIPTDQCNGINRFVALFAVPLLSHFHIS 60
Db 1 MITAADFYHVTAVVPLVYAMTAYGSRVWRRIPTDQCNGINRFVALFAVPLLSHFHIA 60
QY 61 TNDPFAANLRFADLTQKQAVLALLASRGLSSPRALGLDWSITLFSLSSTPNTLVNG 120
Db 61 ANNPYANLRFADLSQKQVIVLSLFLWCK-LS--RNGSLDWTITLFSLSSTPNTLVNG 117
QY 121 IPLLKMGYGAASSAGTLMVQVVLQCIITWYIMLFLFYRAARALVDQPDGAAASIVSF 180
Db 118 IPLLKMGYV-NFSGDLNVQIVLQCIITWYIMLFLFYRAARALVDQPDGAAASIVSI 175

```

```

QY 181 RVSDSVVSL-ARGDVELEAPDGVAGAGAVSRGGDAGRVRVTVRKSTSRSEAAASHSH 239
Db 176 HVSDIMSLDGRQPLETEAIK-----EDGKLHVTVRRSNASRSDIYSRRSQ 222
QY 240 SQTMQPRVSNLSGVLYSLOSSENTPRGSSFNHADFNIV----- 280
Db 223 GLSATPRPSNLTWAEIYSQSRRNPTRGSSFNHTDYSMASGGGNSNFGPGEAVFGS 282
QY 281 -----GAAAK--GGGGAAG--DDEKACCGGGGGHSPQP----- 310
Db 283 KGTPPRPSNYEEDGPAKPTAAGTAAGAGRFHYQSGSGGGGGAHYPAHPGMFSPNTGG 342
QY 311 -----QAVAVPAKR-----KOLHMLWSSSPVSEAAVHVFGAGDADVLAKG 357
Db 343 GGGTAAGKNAVPVGGKQDGNHLMFVWSSASPVSD-----VFGGGGNHH--ADY 394
QY 358 AQAYDEYGRD-----DYSRTKNGSGGADKGPTLSKLG-----SNSTAQLYP 400
Db 395 STATNDHQKDVKISVPCGNSNDNQYVEREEFSGNKDDSKVLATDGNNSNKTQ-- 451
QY 401 KDDGGRAAAVAPPASVMTLILIMVWRKLIENPNYSSLIQVWLSVSYRWGTEMPAI 460
Db 452 -----AKVMPPTSVMTLILIMVWRKLIENPNYSSLIQVWLSVSYRWGTEMPAI 502
QY 461 IARSISILSDAGLGMAFMFSLGLFMALQPRIIACGNKLAIAIAMGVRFVAGPAAVMAASIAV 520
Db 503 IAKSISILSDAGLGMAFMFSLGLFMALNPRIIACGNRRAPFAAAMRFVVGPAVMAVAV 562
QY 521 GLRGVLLHIAIQAALPQGIQVFPVFAKEYGVHDPDILSTA--YG-----PIT 564
Db 563 GLRGVLLHIAIQAALPQGIQVFPVFAKEYGVHDPDILSTGVIFQMLVALPIT 613

RESULT 6
Q9FVF6 PRELIMINARY; PRT; 614 AA.
AC Q9FVF6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE PIN1-like auxin transport protein.
GN PPL1.
OS Populus tremula x Populus tremuloides.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid I; Malpighiales; Salicaceae; Salicaceae; Populus.
OX NCBI_TaxID=47664;
RN [1]
RP SEQUENCE FROM N.A.
RC Schrader J., Bhale Rao R.P., Palme K., Sandberg G.;
RA "The PPL family of PIN1-like auxin transporters in Hybrid Aspen."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF190881; AAG17172.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; I.
DR TIGRFAMs; TIGR00946; 2a69; 1.
SQ SEQUENCE 614 AA; 67254 MW; 7DA7BED27570ACC9 CRC64;

Query Match 55.5%; Score 1614.5; DB 10; Length 614;
Best Local Similarity 55.6%; Pred. No. 9.9e-103;
Matches 362; Conservative 62; Mismatches 94; Indels 133; Gaps 22;

QY 1 MITALDLYHVTAVVPLVYAMTAYGSRVWRRIPTDQCNGINRFVALFAVPLLSHFHIS 60
Db 1 MISTDFYHVTAVVPLVYAMTAYGSRVWRRIPTDQCNGINRFVALFAVPLLSHFHIS 60
QY 61 TNDPFAANLRFADLTQKQAVLALLASRGLSSPRALGLDWSITLFSLSSTPNT 116
Db 61 TNDPFAANLRFADLSQKQVIVLSLFLWCK-LS--RNGSLDWTITLFSLSSTPNT 113
QY 117 LVNGIPLLRMGYGAASSAGTLMVQVVLQCIITWYIMLFLFYRAARALVDQPDGAAAS 176

```

Db 114 LVNGIFLLKMGVGDYS-GSLMVQVVVLOCIITWYTLMLFLEFYRAARALVLDQFPDGAASIVSF 171  
 QY 177 IYSFRVDSVWSL-ARGDVELEABPCVAGAVSSRGCDACRVVTVYRKSTSSRSEAAC 235  
 Db 172 IYSIHVDSIMSLDGRPLETEA-----AIK-----EDGKLHVTYRKSNASRSDIFS 218  
 QY 236 SHSHS-OTWQPRVSNLSGVEIYSLSSRNPTPRGSENFHADFFENIVGAAAGGGGAG-- 292  
 Db 219 RRSQGLSSTTPRPSNLTADEIYSLSSRNPTPRGSSFNHTDFYSMMAAGRNSNFASDVY 278  
 QY 293 -----DEKGAC-----GGGGGHSPOP----- 310  
 Db 279 GLSASRGPTPRPSNFEEHGGGKPKFRFHYHAPGATHYAPAPNPGMPSPTTAASKGVSAN 338  
 QY 311 --QAVAVPAKR-----KDLHMLVWSSSASPVSERAAVHVFGAGGADHADV--- 353  
 Db 339 ANNAARAAAKKNGQAOQAEDGRDLHFWVWSSSASPVSVDVFGHDYCA--HDLKDVRYA 396  
 QY 354 LAKG-----AQAYDEYV-----STKNGSGGAKGQPTLSKLSGNSNTAOLYP 400  
 Db 397 VSPGKVEGQRENGEDYNLEERDDFSQNRGLDRERNSHSG-BKGG----- 439  
 QY 401 KDDGEGRAAAAMPASVMTLILIMVWKLIRNPTNTYSSLIGVWVSLVSVRWGIEMPAI 460  
 Db 440 -FDGKPK-----AMPPTSVMTLILIMVWKLIRNPTNTYSSLIGLWSLVSPFNWKNPLI 494  
 QY 461 IAKSISILSDAGLGWAMFSLGLFMALQPRIIACGNKLAALANGVRFVAGPAAVMAAIAV 520  
 Db 495 IAKSISILSDAGLGWAMFSLGLFMALQPRIIACGNKIATFSMAVRFVAGPAAVMAAIAV 554  
 QY 521 GLRGVLLHAIYQAAALPQGIIVPFVFAKEVGHVPDILSTA--YG-----PIT 564  
 Db 555 GIRGTLHAIYQAAALPQGIIVPFVFAKEVGHVPDILSTLSTGVIFGMLIALPIT 605

## RESULT 7

Q8LKH1 PRELIMINARY; PRT; 640 AA.  
 ID Q8LKH1  
 AC Q8LKH1  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE PIN1-like auxin transport protein.  
 GN PIN3.  
 OS Populus tremula x Populus tremuloides.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.  
 OC NCBI\_TaxID=47664;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Baba K., Schrader J., Palme K., Bhalarao R.P., Sandberg G.;  
 RT "A family of PIN1 like auxin transporters in hybrid aspen.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF515434; AAM54033.1; -  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR InterPro; IPR004776; Auxin\_eff.  
 DR Pfam; PF03547; Auxin\_eff; I.  
 DR TIGRFAMs; TIGR00946; 2a69; I.  
 SQ SEQUENCE 640 AA; 69528 MW; 71C8F4C7C9D40D40 CRC64;

Query Match 55.2%; Score 1605.5; DB 10; Length 640;  
 Best Local Similarity 54.7%; Pred. No. 4; 4e-102;  
 Matches 359; Conservative 71; Mismatches 109; Indels 117; Gaps 19;

QY 1 MITALDLVHLTVAVVPLVYAMTAYGSRVWRWIFTPDQCSGINRFVALFAVPLLSHFHIS 60  
 Db 1 MISWNDLYNVLGSAVPLVYAMTAYGSRVWRWKIFSPDQCSGINRFVAFVPLLSHFHIS 60  
 QY 61 TNDPFAANLRFADDTLQKVAVALALLASRGLSPRALGDSITLFSLSLTPNTLVMG 120  
 Db 61 TNDPYAMNFRFAADTLQKIIMLGIWNTFKNG---SLEWMITIFSSTLTPNTLVMG 117

QY 121 IPLLGMYGASSAGTLMVQVVVLOCIITWYTLMLFLEFYRAARALVLDQFPDGAASIVSF 180  
 Db 118 IPLLTAATGTYS-GSLMVQIIVVLOCIITWYTLMLFLEFYRGAKMLIMEQFPE-TAASIVSF 175  
 QY 181 RYDSDVWSL-ARGDVELEABPCVAGAVSSRGCDACRVVTVYRKSTSSRSEAACSHSH 239  
 Db 176 KYDSDVWSLDRDPLETDAEI-----GDGKLFVTVYRKSNASR--RSLGPGS 220  
 QY 240 SQTMPRVSNLSGVEIYSLSSRNPTPRGSENFHADFFNIVGA----- 282  
 Db 221 FSGMTPRPSNTJGABIIYSLSSRNPTPRGSNFNPSDFYSMMGVQGFGRHSLNPADLYS 280  
 QY 283 --AAKGGGAAAGDEKAGCGGGGHSPO-----POAVAVP-----AKR 319  
 Db 281 VQSSRGPTPRPSNFEE--CAPATLSSPRGFGYPAQVPTSYAPNPEFASVTVTAKN 339  
 QY 320 -----KDLHMLVWSSSASPVSERAAVHVFGA---GGADHADVLAKGA----- 358  
 Db 340 QQQQNSKANHDAKELHFMFWSSSASPVSSEGGLHVFGGTDFGASBOSGRSDQAKETRL 399  
 QY 359 -----QAYDEYGRDDYSRTKNGSGAD---KGGPT-LSKLSGNSNTAOLYP 400  
 Db 400 VADHPQNGETKTIPOQDGDPAGEDFSFAGRGEGDDVDQREKEGPTGLNKLSSSTAEQLP 459  
 QY 401 K-----DDEGEGRAAVAMPASVMTLILIMVWKLIRNPTNTYSSLIGVWVSLVSRWGI 455  
 Db 460 KAAEAPDSGGR-----KMPASVMTLILIMVWKLIRNPTNTYSSLIGLWSLVAFRWHV 515  
 QY 456 EMPATIAISILSDAGLGWAMFSLGLFMALQPRIIACGNKLAALANGVRFVAGPAAVMAA 515  
 Db 516 EMPKIIKQISILSDAGLGWAMFSLGLFMALQPKLIACGNSVATFAMAVRFLTGPAVMAA 575  
 QY 516 ASIAGVGLGHVLLHAIYQAAALPQGIIVPFVFAKEVGHVPDILSTA--YG-----PIT 564  
 Db 576 ASIAGVGLRGLTLHAIYQAAALPQGIIVPFVFAKEVGHVPDILSTAVIFGMLIALPIT 631

## RESULT 8

Q7X9P6 PRELIMINARY; PRT; 619 AA.  
 ID Q7X9P6  
 AC Q7X9P6  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE PIN1-like protein.  
 OS Populus tomentosa.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.  
 OC NCBI\_TaxID=118781;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fan J.H.;  
 RT "Isolation and expression of Populus tomentosa pin1-like gene.";  
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY302060; AAP59843.1; -  
 SQ SEQUENCE 619 AA; 67192 MW; 225F129221B3E9D1 CRC64;

Query Match 54.9%; Score 1599; DB 10; Length 619;  
 Best Local Similarity 53.6%; Pred. No. 1; 2e-101;  
 Matches 354; Conservative 64; Mismatches 100; Indels 142; Gaps 16;

QY 1 MITALDLVHLTVAVVPLVYAMTAYGSRVWRWIFTPDQCSGINRFVALFAVPLLSHFHIS 60  
 Db 1 MICWNDLYNVLGSAVPLVYAMTAYGSRVWRWKIFSPDQCSGINRFVAFVPLLSHFHIS 60  
 QY 61 TNDPFAANLRFADDTLQKVAVALALLASRGLSPRALGDSITLFSLSLTPNTLVMG 120  
 Db 61 TNDPYAMNFRFAADTLQKIIMLGIWNTFKNG---SLEWMITIFSSTLTPNTLVMG 117  
 QY 121 IPLLGMYGASSAGTLMVQVVVLOCIITWYTLMLFLEFYRAARALVLDQFPDGAASIVSF 180

```

Db 118 IPLLITANYGYS--GSLMVQIVLVQCIITWYLLFLFEYRGAKLMEQFPE--TAASIVSF 175
QY 181 RVSDVVVSL--ARGDVELEABPDGVAGAGVSSRGDAGVRVTVKSTSSRSEACSHSH 239
Db 176 KVDSDVVVSLGRDFLETDABI-----GDDGKLHVTVRKSNASR--RSLGPGS 220
QY 240 SQTMPRVSNLGVSEIYSLQSSRNPTPRGSSFNHADPFNIVGAAAGGGAAGDEKAC 299
Db 221 FSGWTPRPSNLTAIEIYSLSSRNPTPRGSNPNFSPDFYSMMGVQ----- 264
QY 300 GGGGGHIS-----POPO-----AVAVPAK----- 318
Db 265 -GPGRHSNLGADLYSVQSSRGTPRPSNPFENECAPTATLSPRFGFYPAQTVPSTYPA 323
QY 319 -----RKDLHMLVWSSSASPVSEAAVHVFGA-----CGA 348
Db 324 PNPEFASTVTTKAKNQOQONSKANHDAXELHMFVWSSSASPVSEGGGLHVFEGTDFGAS 383
QY 349 DHADVLAKGA-----QAYDEYGRDDYSRTKNGSGGAD-----KGGP 385
Db 384 EQGRSDQGAKEIRMLVADHPRNGETKTIPQQDGFAGEDFAGRGEGDDVDQREKEGP 443
QY 386 T-LSKLGSNSTAOIYK-----DDGEGRAAAVAMPASVWTRILILIMVWRKLIKNPTYS 439
Db 444 TGLKLGSSSTAELOPKAAEAPDSGSR-----KMPASVWTRILILIMVWRKLIKNPTYS 499
QY 440 SLGVVWVSLVSWANGTEMPAIIARSISILSDAGLGAMPSLGLFVALOPRIIACGNKLA 499
Db 500 SLIGLWVSLVAFRHWEMPKIHKQISISILSDAGLGAMPSLGLFVALOPKLIACGNSVAT 559
QY 500 IAMGVREVPAGPAAVAAASITAVGLRGVLLHIAIVQAALPOGIVPFFVFAKEYGVHPDILSTA 559
Db 560 FMAVRELTGPAVAAASITAVGLRGVLLHIAIVQAALPOGIVPFFVFAKEYGVHPDILSTA 619

RESULT 9
Q8GV74 PRELIMINARY; PRT; 604 AA.
ID Q8GV74 Q8GV76 Q8GV76 PRELIMINARY; PRT; 659 AA.
AC Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Auxin efflux carrier protein.
GN PIN1.
OS Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
CX NCBI_TaxID=3880;
RN [1]
RP SEQUENCE FROM N.A.
RA Schnabel E.L., Frugoli J.A.;
RT "Putative auxin import and export carrier proteins of Medicago truncatula."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY115839; C: integral to membrane; IEA.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; 1.
DR TIGRFAMs; TIGR00946; 2a69; 1.
SQ SEQUENCE 604 AA; 66491 MW; 3E7AB3499F424676 CRC64;

Query Match 54.9%; Score 1597.5; DB 10; Length 604;
Best Local Similarity 55.2%; Pred. No. 1.4e-101;
Matches 354; Conservative 66; Mismatches 98; Indels 123; Gaps 19;

QY 1 MITALDLYHVLTAIVPLVAMTAYGVSRWRIPTPDQCSGNGRFAVLPVLLSFHIFIS 60
Db 1 MILTDFYHVTAVVPLVAMTAYGVSRWRIPTPDQCSGNGRFAVLPVLLSFHIFIA 60
QY 61 TNDPFAMNLFRLAADTLQKVAIALLAL-----ASRGLSPRALGIDNSITLFSITLPT 116

```

```

Db 61 SNNPYKONIRFLAADTLQKLIILCLLATWSNFSKRGK-----LEWTTITLFSITLPT 113
QY 117 LVMGIPLLAGMYGASSAGTLMVQVVLQCIITWYLLFLFEYRAARALVLDQFPDGAAS 176
Db 114 LVMGIPLLAGMYGEPFS--GSLMVQIVLVQCIITWYLLFLFEYRGARLLISBQFPD--TAGS 171
QY 177 IYSEFRVDSDDVSL--ARGDVELEABPDGVAGAGVSSRGDAGVRVTVKSTSSRSEAAC 235
Db 172 IYSEIHVDSVMSLDGRTPLETDABEK-----EDGKLHIYVRKSNASRSDIYS 218
QY 236 SHSHS--QTMQPRVSNLGVSEIYSLQSSRNPTPRGSSFNHADPFNIVGA--AAKGGGGAAG 292
Db 219 RRSQGLSSNTPRPSNLTAIEIYSLQSSRNPTPRGSSFNHTDFYSMMGGGNGRNSNFAND 278
QY 293 DEEKACACGGGGCHSPQP-----QAVA 314
Db 279 VVNYGLSANSRGVTPRPSNPFENECAPTATLSPRFGFYPAQTVPSTYPA 338
QY 315 VPAKR-----KDLHMLVWSSSASPVSEAAVHVFGAGGADHADV-----LAKG----- 357
Db 339 VNAKKSNGSQOQKQBDLHMFVWSSSASPVSDVFGHDFGA--HDQKEVTLNVSEKVEGH 396
QY 358 AQAYDEY--GRDDYS-----SRTKNGSGCADKGGPTLSKLGSNSTAOIYKPDGEGRAAA 410
Db 397 RETQDYLEKDEFSFGNGKMERENQHGGEGKG-----DGKSK----- 435
QY 411 VAMPASVWTRILILIMVWRKLIKNPTYSILGVVWVSLVSWANGTEMPAIIARSISILSD 470
Db 436 -YMPASVWTRILILIMVWRKLIKNPTYSILGVVWVSLVSWANGTEMPAIIARSISILSD 494
QY 471 AGLGMAFSLGLFVALOPKLIACGNSIAAFSGMVRFLVGPVAVMAAASFAVGLKGVLPFVA 530
Db 495 AGLGMAFSLGLFVALOPKLIACGNSIAAFSGMVRFLVGPVAVMAAASFAVGLKGVLPFVA 554
QY 531 IVQAALPOGIVPFFVFAKEYGVHPDILSTA--YG-----PIT 564
Db 555 IVQAALPOGIVPFFVFAKEYGVHPDILSTGVIFGMLIALPIT 595

RESULT 10
Q8GV76 PRELIMINARY; PRT; 659 AA.
ID Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76
AC Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76 Q8GV76
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Auxin efflux carrier protein.
GN PIN1.
OS Medicago truncatula (Barrel medic).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.
CX NCBI_TaxID=3880;
RN [1]
RP SEQUENCE FROM N.A.
RA Schnabel E.L., Frugoli J.A.;
RT "Putative auxin import and export carrier proteins of Medicago truncatula."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY115839; C: integral to membrane; IEA.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR InterPro; IPR004776; Auxin_eff.
DR Pfam; PF03547; Auxin_eff; 1.
DR TIGRFAMs; TIGR00946; 2a69; 1.
SQ SEQUENCE 659 AA; 71974 MW; F641D1567D40797E CRC64;

Query Match 54.8%; Score 1594; DB 10; Length 659;
Best Local Similarity 52.2%; Pred. No. 2.8e-101;
Matches 360; Conservative 61; Mismatches 103; Indels 166; Gaps 16;

QY 1 MITALDLYHVLTAIVPLVAMTAYGVSRWRIPTPDQCSGNGRFAVLPVLLSFHIFIS 60
Db 1 MITWHDLYTLTAIVPLVAMTAYGVSRWRIPTPDQCSGNGRFAVLPVLLSFHIFIS 60

```

61 TNDPFAANLRLAADTLQKAVLALLALASRLGLSSPRALGDWSITLFLSLTPNLTVMG 120  
 61 TNPYAMWYFLAADTLQKIMLLALTIWTFANG---SLEWMIITLFLSLTPNLTVMG 117  
 121 IPLLKMGYASSAGTLMVQVVLQCIITWYIMLFLFEYRAARALVDQPPDGAASIVSF 180  
 118 IPLLAMYGEYS-GTLMVQVVLQCIITWYIMLFLFEYRAARALVDQPPDGAASIVSF 175  
 181 RVSDVVL-ARGDVELEAEAPDGVAGAGVSSRGDAGRVRVTRKSTSRSEACSHH 239  
 176 KVSDVVLQGRDPLEDAEV-----GDDGKLHVHVRKSNASR-----R 214  
 240 SOTMQRVSNLGSVEIYSLSSRNPTPRGSSFNHADFFNIVGAAAGGGGAAGDEKGCAC 299  
 215 SFWMTPRPSNLTGAEIYSLSSRNPTPRGSGFNHNTDFYSMMGVAPRHSNGANDVYSVQS 274  
 300 GGGGGGHP----- 308  
 275 NSRGTPRPSNFEENPVPTATSAQTVNSPKGFPYPAQNVPAAPNPEPSSGLSKSVSK 334  
 309 -----QPOAV-----AVPAK-----RKDLHMLVWSSASPVSEAAVHV 342  
 335 NSQHLSSQQTQVQVQVQVTSNGSAAAKTSHDAKELHMFVWSSASPVSEASGLQV 394  
 343 FGAGGADHADVLAKGAQAYDEYGRDDYSSR-----TKNGSGADKGG- 384  
 395 FGSGAAYG-----ASDQSGRSEQGAKEIRMLVSDDDHPNGVNIITKMGQETELGGE 446  
 385 -----PTLSKL-GSNSTAQLYPKDDGEGRAAAVAMPASVMT 421  
 447 ELKFPVKEEBLLVVEEVGKVEPAALNKLGGSSSTAELHP-----YATVKEMPPASVMT 500  
 422 LILIMWRKLRNPNTYSSILGVVWSLVSRWNGIEMPAITARSISILSDAGLGNAMFSLG 481  
 501 LILIMWRKLRNPNTYSSILGLIWSIVAFRWHVMEPKIIEKSTISILSDAGLGNAMFSLG 560  
 482 LFMALQRIITACGNKLAIAIANGVRFVAGPAPVMAAASIAVGLRGLVHLIAIVQALPGQIV 541  
 561 LFMALQRIITACGNVSAFNAVAVRFLTGPAVMAAASIAVGLRGLVHLIAIVQALPGQIV 620  
 542 PFVFAKEYGVHPDILSTA--YG-----PIT 564  
 621 PFVFAKEYGVHPDILSTGVIFGMLIALPIT 650

## RESULT 11

Q84WX3 ID Q84WX3 PRELIMINARY; PRT; 599 AA.  
 AC Q84WX3  
 DT 01-JUN-2003 (TREMELrel. 24, Created)  
 DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Auxin efflux carrier protein PIN1.  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Viciae; Pisum.  
 OC NCBI\_TaxID=3888;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Chawla R., DeMason D.A.;  
 RT "PSPIN1, putative auxin efflux protein.";  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AV222857; AAC38045.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR InterPro; IPR004776; Auxin\_eff.  
 DR Pfam; PF03547; Auxin\_eff; 1.  
 DR TIGRFAMs; TIGR00946; 2a69; 1.  
 SQ SEQUENCE 599 AA; 66455 MW; A835C7759FEFA2B CRC64;

Query Match 54.5%; Score 1587; DB 10; Length 599;  
 Best Local Similarity 56.2%; Pred. No. 7.5e-101;

Matches 351; Conservative 70; Mismatches 108; Indels 96; Gaps 18;  
 QY 1 MITADLVHVLTAVVPLVAMTAYGSRVWRRIPTDQCSGINRFVALFAYVLLSFHFIS 60  
 Db 1 MITLDFHVTAMVPLVAMTAYGSRVWRRIPTDQCSGINRFVALFAYVLLSFHFIA 60  
 QY 61 TNDPFAANLRLAADTLQKAVLALLAL-----ASRGLSSPRALGDWSITLFLSLTPNLT 116  
 Db 61 SNNPYKMLRLAADTLQKIMLLTLFTWSNFSKRG-----SLEWITLFLSLTPNLT 113  
 QY 117 LVMGPIPLKMGYDGS-GSLMVQIVVLCIITWYIMLFLFEYRAARALVLDQPPDGAAS 176  
 Db 114 LVMGPIPLKMGYDGS-GSLMVQIVVLCIITWYIMLFLFEYRAARALVLDQPPDGAAS 171  
 QY 177 IVSFRVSDVVL-ARGDVELEAEAPDGVAGAGVSSRGDAGRVRVTRKSTSRSEAC 235  
 Db 172 IVSIVHSDVMSLDGRTPLETDAEIK-----QDGKLVTVTRKSNASRSDIYS 218  
 QY 236 SHSHS-QTMQPRVSNLGSVEIYSLQSSRNPTPRGSSFNHADFFNIVG----- 281  
 Db 219 RRSQGLSNTFRPSNLTAIEIYSLQSSRNPTPRGSSFNHTDFYSMMGGGRNSNFNADVN 278  
 QY 282 -----AAAKGGGAAGDEKGCACGGGGGSHSPQO-AVAVPA----- 317  
 Db 279 NYGLSASRGVTPRPSNYEEDASNAKLYHPAPNPGMFSPNTKNLGSNNVVKRSGNQD 338  
 QY 318 -----KRDILHMLVWSSASPVSEAAVHVFGAGADHADV---LAKGAQAYDEYGRDD 368  
 Db 339 QNQNQKQDDLHMFVWSSASPVSDVFGGHEFGS--HDQKEVKLVNPKVDGHRTEQED 396  
 QY 369 YSRRTKNGSGADKGGPTLSKLGNSSTAQLYPKDDGE--GRAAAVAMPASVMTLILIM 426  
 Db 397 YLEKDEFSFG--NRG---MEREMNQ-----QHEGEKIGDGKSKVMPASVMTLILIM 445  
 QY 427 WRKLRNPNTYSSILGVVWSLVSRWNGIEMPAITARSISILSDAGLGNAMFSLGLFWAL 486  
 Db 446 WRKLRNPNTYSSILGLIWSIVAFRWHVMEPKIIEKSTISILSDAGLGNAMFSLGLFWAL 505  
 QY 487 QPRTIACGNKLAIAIANGVRFVAGPAPVMAAASIAVGLRGLVHLIAIVQALPGQIVPVFA 546  
 Db 506 QPRTIACGNKLAIAIANGVRFVAGPAPVMAAASIAVGLRGLVHLIAIVQALPGQIVPVFA 565  
 QY 547 KEYGVHPDILSTA--YG-----PIT 564  
 Db 566 KEYGVHPDILSTGVIFGMLIALPIT 590

## RESULT 12

Q8GV75 ID Q8GV75 PRELIMINARY; PRT; 621 AA.  
 AC Q8GV75;  
 DT 01-MAR-2003 (TREMELrel. 23, Created)  
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)  
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)  
 DE Auxin efflux carrier protein.  
 OS Medicago truncatula (Barrel medic).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.  
 OC NCBI\_TaxID=3880;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RA Schnabel E.L., Frugoli J.A.;  
 RT "Putative auxin import and export carrier proteins of Medicago truncatula.";  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AV115838; AAM55299.1; -;  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR InterPro; IPR004776; Auxin\_eff.  
 DR Pfam; PF03547; Auxin\_eff; 1.  
 DR TIGRFAMs; TIGR00946; 2a69; 1.  
 SQ SEQUENCE 621 AA; 68271 MW; 05DB66F8CF842A8A CRC64;

Query Match 54.4%; Score 1584; DB 10; Length 621;  
Best Local Similarity 54.9%; Pred. No. 1.3e-100;  
Matches 355; Conservative 63; Mismatches 111; Indels 118; Gaps 17;

QY 1 MITALDYLHVLTAIVPLVAVMTLAYGSRVWRRIPTDQCGINRFVALFAVPLLSPHFIS 60  
DB 1 MITLKDLTVLTAVVPLVAVMTLAYGSRVWRRIPTDQCGINRFVALFAVPLLSPHFIS 60  
QY 61 TNDPAMNRLFLAATDLOKVAVALALAL-ASGLSSPRALGLDMSITLFSLSLTPNT 116  
DB 61 TNNPYQMMFRFAATDLOKIMLVALSLSLWTFKNG-----NLEWMTIFLSLTPNT 113  
QY 117 LVMGIPLLRGMYGASSAGTLMVQVVLQCIITWYTLMLFLFYRAARALVDPFGGAAS 176  
DB 114 LVMGIPLLIAMYGDS-GSLMVQVVLQCIITWYTLMLFLFYRAARALVDPFGGAAS 171  
QY 177 IVSFVDSVSL-ARGDVELEAEADPDGAGAGVSSRGDAGRVVTRKTSRSESAAC 235  
DB 172 IVSFVDSVSLDGRDLETDASV-----GDDGKLHVTVRKSNAASRRSFM 218  
QY 236 SHSHSQTPQPRVSNLSGVEIYSLQSRNPTPRGSGFNHADFNTVGAAGGGGAAGD-- 293  
DB 219 N-----TPRPSNLTAETIYSLSS-----TPRGSNFEQEFYSNMGYQPRHSNFGANDLY 267  
QY 294 -----EKGACGGGGGGH----- 306  
DB 268 SVQSSSRGPTPRPSNFEENGASPRFGYPAAQTVPSTYPVNPFEFSSTTKPVKNQNL 327  
QY 307 ---SPQQAIVAPAKKDLHMLVWSSSPVSEAAHVH-----GAGGADHADVLA-- 355  
DB 328 MPQQFQVQLTKGSDKELHMFVWSSSPVSEAGLVNFSBQSEBAGEKIRVVVADE 387  
QY 356 -----KGAQAYDEYGRDYS-SRTKNGSGGADKGGPTLSKLSNSTAQLPKDGG- 404  
DB 388 HNQNGETNKGQKEIGEEDFKNGVKVGEQEGPNP--NKLGSNSTPELHPKATGV 445  
QY 405 EGRAAIVAMPASVMTLILIMVWKLIRNPNTYSSLLGVVMSLVSVRWGTEMPAIIARS 464  
DB 446 ADSVGVKLMPPASVMTLILIMVWKLIRNPNTYSSLLGLVSLVAFVDFVEMPKIEKS 505  
QY 465 ISILSDAGLGMAFSLGLFMALQPIRIACGNKLAIAIMGVFVAGPAPVAAAASAVGLRG 524  
DB 506 ISILSDAGLGMAFSLGLFMALQPKIACGNVSAFAMAIPLTGPVAAAASAVGLRG 565  
QY 525 VLLHIAVQALPQGIIVPFAKEYGVPDILSTA--YG-----PIT 564  
DB 566 TLLHIAVQALPQGIIVPFAKEYNVHPAILLSTAVIFGLIALPIT 612

RESULT 13  
Q987Z8 PRELIMINARY; PRT; 640 AA.

AC Q987Z8; 13, Created)  
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TEMBLrel. 24, Last annotation update)  
DE Auxin transport protein (Putative auxin transport protein RBH1).  
GN F15H11.14 OR PIN3 OR AT1G70940.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
OC NCBI\_TaxID=3702;  
RN [1]  
SEQUENCE FROM N.A.  
RA Federapfel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.P.,  
RA Altari H., Araujo R., Huizar L., Rowley D., Buehler E., Dunn P.,  
RA Gonzalez A., Kremenetskaia I., Kim C., Lenz C., Li J., Liu S.,  
RA Lueros S., Schwartz J., Shinn P., Toriumi M., Vysotskaia V.S.,  
RA Walker M., Yu G., Ecker J., Theologis A., Davis R.W.,  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.

SEQUENCE FROM N.A.  
RP Primi J., Wisniewska J., Palme K.,  
RT "PIN gene family in Arabidopsis thaliana";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.

[3]  
RN  
SEQUENCE FROM N.A.  
RP Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,  
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,  
RA Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,  
RA Kawai J., Kwei C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,  
RA Saki M., Shinn P., Tang C.C., Toroumi M., Wallender E.K., Wong C.,  
RA Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,  
RA Theologis A., Davis R.W.,  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

[4]  
RN  
SEQUENCE FROM N.A.  
RP Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,  
RA Palm C.J., Bower L., Jones T., Banh J., Carninci P., Chen H.,  
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,  
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,  
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,  
RA Ecker J., Theologis A., Davis R.W.,  
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC008148; AAD55507.1; -  
DR EMBL; AF087818; AAD52695.1; -  
DR EMBL; AF136327; AAM96993.1; -  
DR EMBL; BT002085; AAN72096.1; -  
DR PIR; G96733; G96733.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR InterPro; IPR004776; Auxin\_eff.  
DR Pfam; PFO3547; Auxin\_eff; 1.  
DR TIGRFAMs; TIGR00946; 2a69; 1.  
SEQUENCE 640 AA; 69465 MW; 256F82C8E1ADADB0 CRC64;

Query Match 53.6%; Score 1560.5; DB 10; Length 640;  
Best Local Similarity 52.8%; Pred. No. 5.4e-99;  
Matches 353; Conservative 73; Mismatches 102; Indels 139; Gaps 20;

QY 1 MITALDYLHVLTAIVPLVAVMTLAYGSRVWRRIPTDQCGINRFVALFAVPLLSPHFIS 60  
DB 1 MISHDLYTLTAIVPLVAVMTLAYGSRVWRRIPTDQCGINRFVALFAVPLLSPHFIS 60  
QY 61 TNDPAMNRLFLAATDLOKVAVALALALASRLGSLSPRALGLDMSITLFSLSLTPNTLVMG 120  
DB 61 TNNPYAMNRLFLAATDLOKIMLSLVLA--NFTRSGLSLEWSTIFSLSLTPNTLVMG 117  
QY 121 IPLLGMYGASSAGTLMVQVVLQCIITWYTLMLFLFYRAARALVDPFGGAASIVSF 180  
DB 118 IPLLIMYGEYS-GSLMVQIVLQCIITWYTLMLFLFYRAARALVDPFGGAASIVSF 175  
QY 181 RVSDVSVSLARGD-VELEAPDPGAVAGVSSRGDAGRVVTRKTSRSENAACSHSH 239  
DB 176 KVESDVSVSLDGHDFLETDAL- -GDDGKLHVTVRKSNAAS-RSFC- - - 217  
QY 240 SQTMQPRVSNLSGVEIYSLQSRNPTPRGSGFNHADFNTVGAAGGGGAAGD--A 282  
DB 218 GPNTFPSPNLTAETIYSLST-----TPRGSNFEHSDFYNMGMFGGRLSNFGPADMYSVQ 273  
QY 283 AAKG-----GGGAAGDEKAGCGGGGGHSPQ----- 310  
DB 274 SSRGTPRPSNFEENGASPRFGYPGGGAGSYAPNPPEFSSTTTSTANKSVNKNPKD 333  
QY 311 ----QAVAVPA-----KKDLHMLVWSSSPVSEAAHVHFGAGGADH-----ADVL 354  
DB 334 VTNQQTTLFTGKSKNSHDAKELHMFVWSSNGSPVSDRAGLVFG-GAPDNDQGRSDQG 392  
QY 355 AK-----GAQAYDEYGRDDYSRRTKNGSGGADKGGPTL 387  
DB 393 AKETRLMLVPDQSHNGETKVAHHPASGDFGEGEQQFSFAGKEEBAERPKDAENG-----L 445  
QY 388 SKLGSNSTAQLPKDGGEGAAVA--MPPASVMTLILIMVWKLIRNPNTYSSLIGV 444  
DB 446 NKLAPNSTAALQSK-TGLGGAEASQRKNMPPASVMTLILIMVWKLIRNPNTYSSLIGL 504



Search completed: March 3, 2004, 09:08:44  
Job time : 75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:05:58 ; Search time 23 Seconds  
(without alignments)  
1286.160 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910

Sequence: 1 MITALDLYHVLRAVPLVLA.....DILSTAYGPTTSHGFITCHS 573

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*
- 2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*
- 3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep.\*
- 4: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*
- 5: /cgn2\_6/prodata/2/iaa/PCTUS\_COMB.pep.\*
- 6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	144	4.9	320	4	US-09-489-039A-13804
2	138.5	4.8	738	3	US-08-864-038A-3
3	132	4.5	3546	4	US-09-679-279-13
4	126.5	4.3	984	4	US-09-612-402B-43
5	126.5	4.3	1012	4	US-09-612-402B-2
6	122.5	4.2	982	4	US-09-556-877-176
7	122.5	4.2	982	4	US-09-620-412C-176
8	122.5	4.2	982	4	US-09-598-419-176
9	122.5	4.2	1006	4	US-09-556-877-190
10	122.5	4.2	1006	4	US-09-620-412C-190
11	122.5	4.2	1006	4	US-09-598-419-190
12	122.5	4.2	1013	4	US-09-612-402B-16
13	121	4.2	505	4	US-09-612-402B-17
14	121	4.2	1501	4	US-09-252-991A-20266
15	119.5	4.1	542	4	US-09-252-991A-26256
16	118	4.0	503	4	US-09-252-991A-32643
17	116.5	4.0	509	4	US-09-252-991A-23576
18	114.5	3.9	582	4	US-09-450-072-81
19	114.5	3.9	582	4	US-09-351-348-81
20	114.5	3.9	1482	4	US-09-410-551B-21
21	114.5	3.9	1488	4	US-09-410-551B-17
22	113.5	3.9	553	4	US-09-252-991A-17984
23	113.5	3.9	729	4	US-09-291-922-29
24	113.5	3.9	1013	4	US-09-612-402B-15
25	113	3.9	613	4	US-09-489-039A-13221
26	113	3.9	1700	4	US-09-252-991A-21763
27	111.5	3.8	302	4	US-09-252-991A-29977

```

28 110.5 3.8 383 4 US-09-252-991A-24300
29 110.5 3.8 748 2 US-08-997-080-154
30 110.5 3.8 748 2 US-08-997-362-154
31 110.5 3.8 748 3 US-09-095-855-154
32 110.5 3.8 748 4 US-09-324-542-154
33 110.5 3.8 748 4 US-09-205-426-154
34 109 3.7 1248 3 US-08-726-214-16
35 108 3.7 551 4 US-09-489-039A-9083
36 107.5 3.7 421 3 US-09-000-040-1
37 107.5 3.7 742 3 US-08-791-115B-5
38 107 3.7 4928 3 US-09-036-987A-5
39 107 3.7 4928 3 US-09-370-700-5
40 107 3.7 4928 4 US-09-603-207-5
41 106.5 3.7 278 3 US-09-260-527-3
42 106.5 3.7 280 3 US-09-260-527-1
43 106.5 3.7 526 4 US-09-252-991A-28761
44 106 3.6 317 4 US-09-489-039A-9943
45 106 3.6 508 4 US-09-252-991A-17606

```

#### ALIGNMENTS

```

RESULT 1
US-09-489-039A-13804
; Sequence 13804, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13804
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13804

```

```

Query Match 4.9%; Score 144; DB 4; Length 320;
Best Local Similarity 28.3%; Pred No. 2.7e-05;
Matches 43; Conservative 28; Mismatches 63; Indels 18; Gaps 5;

Qy 404 GEGRAAAVMPASVMTKLILIMVWRKLRINPTYSYSLIGVMSLVYRGMPIAAR 463
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 160 GENGSGTILMLP-----VLWRS-VKPIVWGPLLGVLSAI---GKMPDILLA 205
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 464 SISILSDAGLGWMPFSLGLFWALQPIIACGNKLAATAMGVRFVAGPAAVAAAGTAVGLR 523
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 206 SKPFLGATAAALPLTGVLSARKLQI---NALIATSTIVKLIVQPIAAGLVMLGLH 262
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 524 GVLHIAIVQAALPGQIVFFVFAKEYGVH-PD 554
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 263 GSIAITAILMALAAGFFGVVFGNRFVGQSPD 294
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 2
US-08-864-038A-3
; Sequence 3, Application US/08864038A
; Patent No. 6001592
; GENERAL INFORMATION:
; APPLICANT: Kunio NAKASHIMA et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDE GENE CDNA, VECTOR
; CONTAINING SAID CDNA, HOST CELLS TRANSFORMED WITH SAID
; TITLE OF INVENTION: VECTOR, POLYPEPTIDE PRODUCED THEREBY, METHOD OF PRODUCING
; TITLE OF INVENTION: SAID POLYPEPTIDE, DNA ENCODING SAID POLYPEPTIDE AND ANTIBODY
; TITLE OF INVENTION: TO SAID POLYPEPTIDE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:

```

ADDRESS: 812-5 Hirano  
STREET: Iseshinden  
CITY: Tsu-city  
STATE: Mie-prefecture  
COUNTRY: JAPAN  
ZIP: 514-01  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.50 inch, 1.44 MB storage  
OPERATING SYSTEM: Microsoft Windows 95  
SOFTWARE: Word Perfect 6.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/864,038A  
FILING DATE: May 28, 1997  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 8-194459  
FILING DATE: 15-July-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: C. Bruce Hamburg  
REGISTRATION NUMBER: 22,389  
REFERENCE/DOCKET NUMBER: F-5610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)966-2340  
TELEFAX: (212)953-7733  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 738  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
ORIGINAL SOURCE:  
ORGANISM: Pinctada fucata  
CELL TYPE: mantle epithelial cell  
FEATURE:  
NAME/KEY: peptide  
LOCATION: from 1 to 738  
IDENTIFICATION METHOD: E (by experiment)  
US-08-864-038A-3

Query Match 4.8%; Score 138.5; DB 3; Length 738;  
Best Local Similarity 21.6%; Pred. No. 0.00032;  
Matches 99; Conservative 27; Mismatches 176; Indels 157; Gaps 13;  
QY 202 GVAGAGVSSGGGAGRVTVTRKSTSSRGEACSHSHSQTMQPRVNSLGSVEIYSLOSS 261  
DB 266 GLGGGLGGLGGYGG-----SAAAAAAAAAAGG-----GGLGGVGFYGGGG 310  
QY 262 RNPTPRGSPFNHADFNIIVGAAKGGGGAAGDEKAGCGGGGGHSPQPOAVAVPAKR-- 319  
DB 311 RRGGRGRRRAAAAAAAAAAAGGGGGGGG---GGGGGGGAGAGAAAAAASASASR 367  
QY 320 -----KDLHMLVWSSSASP 333  
DB 368 QMSGJRDALGDIKOLLRSNGASAKASAVASTKSIQIDDLKDLKDLGLLKSASAS 427  
QY 334 USRAAVHVFAGGADH-----ADVLAAGQAYDEYGRDDYSRRTKNGSGGADK 382  
DB 428 ASASASAGGGGGGGGGGGGGGGGGGGGAGALAAAGAGG-----GLGGGGG 477  
QY 383 GCGTTLTKL-----GSNSTAQLYPKDDGEGRAAAVAMPASVMTLILIMVWKLIRN 434  
DB 478 GGAALAAALAAAGAGGGGGLGGLGGGSAASAAAAAAGGGGRALREALRQWEG 537  
QY 435 PNTYSSLIGVWVSLVSTWGTIEMPAITARSISILSDAGLGNAM---FSLGLFWALQPII 491  
DB 538 GGSAAAAAAGGGGGGGG-----GGGFGVGLGGGFGGGF----- 576  
QY 492 ACNKLAAIAVMGRFVAG-----PAYMAAASIAVGLRGLVLIH 529  
DB 577 -GGGSSAAAAAAGAGGGGRRGRGRRGGGDDGNGGASAVAAAAAAGGSAADV 635  
QY 530 AIVQALPQGI VPFVFAKEYGVHDPILSTAYGPI TSHGF 568

DB 636 AAAAAA-----AMYGAGAD-----GPDFDNGF 659  
RESULT 3  
US-09-679-279-13  
Sequence 13, Application US/09679279  
Patent No. 6524841  
GENERAL INFORMATION:  
APPLICANT: McDaniel, Robert  
TITLE OF INVENTION: Recombinant Megalomicin Biosynthetic  
TITLE OF INVENTION: Genes and Uses Thereof  
FILE REFERENCE: 300622004700  
CURRENT APPLICATION NUMBER: US/09/679,279  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/158,305  
PRIOR FILING DATE: 1999-10-08  
PRIOR APPLICATION NUMBER: US 60/190,024  
PRIOR FILING DATE: 2000-03-17  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 3546  
TYPE: PRT  
ORGANISM: Micromonospora megalomicea  
US-09-679-279-13

Query Match 4.5%; Score 132; DB 4; Length 3546;  
Best Local Similarity 22.3%; Pred. No. 0.004;  
Matches 137; Conservative 54; Mismatches 212; Indels 210; Gaps 26;  
QY 1 MITALDYHLTAVVPLVYVAMTAYGVSVMRIETPDQCSGINRFVALFAVPLLSFFHIS 60  
DB 1731 LLATVDDEHPLSAV--FHVAATLDDGTV---ETLTGDIRERANAKVL----- 1773  
QY 61 TNDPPFAMNIRFLAADTLQKVAVLALLASRGLSPRALGLDWSITLFLSLTLPNTLVM- 119  
DB 1774 ---GARNLHETRD-----ADLDAVLFSSSTAAFGAPLG-----GYVPGNAYLD 1816  
QY 120 -----GIFLLRGMYGASSAGTLMVQVVVLOCIIWYLMFLFYRAARALVLDQFP 170  
DB 1817 GLAQQRSEGLPATSVAWG-TWAGSGMAEGPVAD-----RFRHGVMEHP 1861  
QY 171 DGA-----AASIVSFRVDS-----DVVSLARGDVELEAEPDGV 203  
DB 1862 DQAVEGLRVALVQGEVAPIVVDIRWDREPLLAYTAQRTPLFDLDEAR-----RAAPGPD 1916  
QY 204 AGAGAVSRGGDAGRVTVTRKSTSSRSEACSHSHSQTM-----OPRVNLSGVETIY 256  
DB 1917 AGFGVAALAGLEVGEREKAVLDLVRTHAAAVLGHASAEQVDFRAFAELGVDLSLAL- 1975  
QY 257 SLOSSRNPTRG-----SSFNHADFFNIIVGAAKGGGGAAGDEK- 297  
DB 1976 ---RNLTTATGVLATITVFDHPDVRTLAGHLAAELGGSGRERPGGEAPTVAPTDEPI 2032  
QY 298 -----ACGGGGGGHSPQ-----PQAVAVPAKDKLHMLVWSSSASPVSE---RA 338  
DB 2033 AIVGMACRLPGGVDSPEQLWELIVSGRDTASAPGDKS-----WDFALMVSDTTTTRT 2086  
QY 339 AVHVF--GAGGADHA-----DVLAKGAQ-----AYDEYGRDDYSSRTKNGSGG 379  
DB 2087 AFGNFMFGAGEFAAFPGISPREALAMPQQRHALETWEALENAGIRPESLR-----G 2140  
QY 380 ADKGGTILSKNSNTAQLYPKDDGEGRAAAVAMPASVMTLILIMVWKLIRNPNYIS 439  
DB 2141 TDTGVFMGSHGYATGRPKPEDEVDG-----YLTGNTAS 2176  
QY 440 SLIGVVWVSLVSTWGTIEMPAITARSISILSDAGLGNAMFSLGLFWALQPIIACGNKLA 499  
DB 2177 ---VAGRIAYVLGLEGPATVDTACSSSLVALHVAAGSL-----RSGDCG---LA 2221  
QY 500 IANGVRFVAGPAV 512

DB 2222 VAGGVSVWAGPEV 2234

# RESULT 4

US-09-612-402B-43  
 ; Sequence 43, Application US/09612402B  
 ; Patent No. 6642023  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jackson, W. James  
 ; APPLICANT: Pace, John  
 ; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
 ; Patent No. 6642023  
 ; FILE REFERENCE: 7969-086-999  
 ; CURRENT APPLICATION NUMBER: US/09/612,402B  
 ; CURRENT FILING DATE: 2000-07-06  
 ; PRIOR APPLICATION NUMBER: 08/942,596  
 ; PRIOR FILING DATE: 1997-10-02  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SOFTWARE: Patent in version 3.0  
 ; SEQ ID NO 43  
 ; LENGTH: 984  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia sp.  
 ; US-09-612-402B-43

Query Match 4.3%; Score 126.5; DB 4; Length 984;  
 Best Local Similarity 22.3%; Pred. No. 0.0066;  
 Matches 123; Conservative 70; Mismatches 196; Indels 163; Gaps 29;  
 QY 36 PDQCSG--INRFVALFAVPLLSFPHI--STNDPFFAMNLRFLAADTLQKV-----80  
 DB 50 PLSCFGNLLGSFTVLRGHSLTFTENIRTSINGAALSN---SAADGLFTIEGPKELSFNSNC 106  
 QY 81 -AVLALL--ALASRGLSSPRALGLDWSITLFSLSLTPLNTLVWG-----IPLLRGMVGA 130  
 DB 107 NSLLAVLPAATTKGSGTPTTSTPNGHIYSKTDL---LLANNEKFSFYNSLVSGDGA 163  
 QY 131 SSAGTLMVQVVVLCIIWYTLMLFLFEYRAARALVLDQFPDGAASIV-SFRVSDVVSL 189  
 DB 164 IDAKSLTVQGISKLCV-----FQENTAQA-----DGGACQVVTFSFSAANEAPI 207  
 QY 190 ARGDVELEAEAPDGVAGAGAVSSRGDAGRVVTVRKSTSSRSE-AACSHSHSQTMQPRVS 248  
 DB 208 A-----FVANVAGVRGGIAAVDQGGQ-----VSSSTSTEDPVVSPSRNTAVEFDGVA 257  
 QY 249 NLGSGVEIYSIQ-----SSRNPTPRGSSFNHADFNIVGAAAKG 286  
 DB 258 RVGG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGQASNTSNNGDGGAI FCK 315  
 QY 287 GCGAAGDEEKG-----ACGGGGGHSPOQAVAPAKRKDLHMLVWSSS 330  
 DB 316 NGAQAGSNNGSVSFDGEGVVFSSNVAAKGK-----AIYAKKLSV-----AN 359  
 QY 331 ASPV-----SERAHVHFGAG-----GADHADVLAKG-----AQAYDEYGRDDYSR 372  
 DB 360 CGPVQFLRNANDGGAIVLGESGELSADYGDIIFDGNLKRKTAKENAADVNGVTVSSQA 419  
 QY 373 TKNGSGG-----ADKGG-----PTLSKLGNSSTAQ-----LYPKDDGEGRAAAVAMP-PA 416  
 DB 420 ISMGSGGKITTLAKAGHQHLLFNDPIEMANGNQPQOSSKLLKINDGEGYTDIVFANGS 479  
 QY 417 SVMTRELLIMVWRKLIRN-----PNTYSSL-----IGVWVSLVSYRWGIEMPAIARS 464  
 DB 480 STLYQNVITIEQGRIVLREKAKLSVNSLQSGSLYMEAGSTWDFVTPPQPPQPPA--ANQ 537  
 QY 465 ISILSDAGLQMA 476  
 DB 538 LITLSNLHLSLS 549

# RESULT 5

US-09-612-402B-2  
 ; Sequence 2, Application US/09612402B

; Patent No. 6642023  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jackson, W. James  
 ; APPLICANT: Pace, John  
 ; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof  
 ; Patent No. 6642023  
 ; FILE REFERENCE: 7969-086-999  
 ; CURRENT APPLICATION NUMBER: US/09/612,402B  
 ; CURRENT FILING DATE: 2000-07-06  
 ; PRIOR APPLICATION NUMBER: 08/942,596  
 ; PRIOR FILING DATE: 1997-10-02  
 ; NUMBER OF SEQ ID NOS: 43  
 ; SOFTWARE: Patent in version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 1012  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia sp.  
 ; US-09-612-402B-2

Query Match 4.3%; Score 126.5; DB 4; Length 1012;  
 Best Local Similarity 22.3%; Pred. No. 0.0069;  
 Matches 123; Conservative 70; Mismatches 196; Indels 163; Gaps 29;  
 QY 36 PDQCSG--INRFVALFAVPLLSFPHI--STNDPFFAMNLRFLAADTLQKV-----80  
 DB 78 PLSCFGNLLGSFTVLRGHSLTFTENIRTSINGAALSN---SAADGLFTIEGPKELSFNSNC 134  
 QY 81 -AVLALL--ALASRGLSSPRALGLDWSITLFSLSLTPLNTLVWG-----IPLLRGMVGA 130  
 DB 135 NSLLAVLPAATTKGSGTPTTSTPNGHIYSKTDL---LLANNEKFSFYNSLVSGDGA 191  
 QY 131 SSAGTLMVQVVVLCIIWYTLMLFLFEYRAARALVLDQFPDGAASIV-SFRVSDVVSL 189  
 DB 192 IDAKSLTVQGISKLCV-----FQENTAQA-----DGGACQVVTFSFSAANEAPI 235  
 QY 190 ARGDVELEAEAPDGVAGAGAVSSRGDAGRVVTVRKSTSSRSE-AACSHSHSQTMQPRVS 248  
 DB 236 A-----FVANVAGVRGGIAAVDQGGQ-----VSSSTSTEDPVVSPSRNTAVEFDGVA 285  
 QY 249 NLGSGVEIYSIQ-----SSRNPTPRGSSFNHADFNIVGAAAKG 286  
 DB 286 RVGG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGQASNTSNNGDGGAI FCK 343  
 QY 287 GCGAAGDEEKG-----ACGGGGGHSPOQAVAPAKRKDLHMLVWSSS 330  
 DB 344 NGAQAGSNNGSVSFDGEGVVFSSNVAAKGK-----AIYAKKLSV-----AN 387  
 QY 331 ASPV-----SERAHVHFGAG-----GADHADVLAKG-----AQAYDEYGRDDYSR 372  
 DB 388 CGPVQFLRNANDGGAIVLGESGELSADYGDIIFDGNLKRKTAKENAADVNGVTVSSQA 447  
 QY 373 TKNGSGG-----ADKGG-----PTLSKLGNSSTAQ-----LYPKDDGEGRAAAVAMP-PA 416  
 DB 448 ISMGSGGKITTLAKAGHQHLLFNDPIEMANGNQPQOSSKLLKINDGEGYTDIVFANGS 507  
 QY 417 SVMTRELLIMVWRKLIRN-----PNTYSSL-----IGVWVSLVSYRWGIEMPAIARS 464  
 DB 508 STLYQNVITIEQGRIVLREKAKLSVNSLQSGSLYMEAGSTWDFVTPPQPPQPPA--ANQ 565  
 QY 465 ISILSDAGLQMA 476  
 DB 566 LITLSNLHLSLS 577

# RESULT 6

US-09-556-877-176  
 ; Sequence 176, Application US/09556877  
 ; Patent No. 6432916  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Probst, Peter  
 ; APPLICANT: Bhatia, Ajay  
 ; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Fling, Steve

APPLICANT: Maisonneuve, Jeff  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C5  
; CURRENT APPLICATION NUMBER: US/09/556.877  
; CURRENT FILING DATE: 2000-04-19  
; NUMBER OF SEQ ID NOS: 305  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 176  
; LENGTH: 982  
; TYPE: PRT  
; ORGANISM: Chlamydia  
; NAME/KEY: VARIANT  
; LOCATION: (1)---(982)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-556-877-176

Query Match 4.2%; Score 122.5; DB 4; Length 982;  
Best Local Similarity 21.4%; Pred. No. 0.016;  
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;  
QY 36 PDQCSG--INRFVALFAVPLLSFHF--STNDPPAMNLRFLAADTLQKV----- 80  
Db 48 PLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALS---SAADGLFTIEGKELSFNC 104  
QY 81 -AVLALL--ALASRGLSPRALGLDWSITLFSLSLTPNLTVMG-----IPLRGMYGA 130  
Db 105 NSLLAVLPAATTKGSGTPTTSTPSNGTTSYKTDL---LLNNEKFSFVSNLVSQDGA 161  
QY 131 SSAGTLMGVVVLQCIWYLMFLFEYRAARALVLDQPDGAAASIV-SRVDSDVVS 189  
Db 162 IDAKSLTVQGISKLCV-----FOENTAQA-----DGAQCVVTSFAMANEAPI 205  
QY 190 ARGDELEAEPPGVAGAGAVSSRGDAGRVRVTVRKSTSSRSB-AACSHSHSQTWQPRVS 248  
Db 206 A-----FVANVAGVRGGIAAVDQCGQ-----VSSSTSTEDPVVSFRNTAVEFDGNA 255  
QY 249 NLSGVEIYSLQ-----SSRNPTRGSSFNHADFENIVGAAAKG 286  
Db 256 RVGG-GIYSGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGQASNTSNYDGGGAI FCK 313  
QY 287 GGAAGDEBK-----ACGGGGHSPQQA VAVPAKRDHMLVWSSS 330  
Db 314 NGAQAGSNNSGVSFDGEGVVFSSNVAAGKG-----AIYAKKLSV-----AN 357  
QY 331 ASPV-----SRAAVHVFAG-----GADHADVLAKG-----AQAYDEYGRDDYSR 372  
Db 358 CGPVQFLNIANDGGAIIYLGESGELSADYGIIFDGNLKRKTAKENAADVNGVTVSQA 417  
QY 373 TNGSGG-----ADKGG-----PTLSKLSNSTAQ-----LYPKDDGEGRAAAVAMPAS 417  
Db 418 ISMSGGKITTIRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTDIVFANGS 477  
QY 418 VMTRLILIMWRKLI RNPNTYSSLGIVVMSLVSYRWGIEMPAIAR-----SISILSDAG 472  
Db 478 -----STLYQNVITIEQGRIVLRKAKLSVNSLSQTG 508  
QY 473 --LGMAMFSLGLFVALQPRIIACGNKLAIAIANGVRFVAGPAVMAAASIAVGLRGVLLHIA 530  
Db 509 GSLYMEAGSTLDFVTQP-----PQPPAANQLITLSN--LHLS 545  
QY 531 IVQAAALPQGI V-PFVFAKEYGVHPDIL-STAYGPTITSHG 567  
Db 546 LSSLLANNVNTNPTNPPAQDSDHPAVIGSTTAGSVTISG 584

## RESULT 7

US-09-620-412C-176  
; Sequence 176, Application US/09620412C  
; Patent No. 6448234  
; GENERAL INFORMATION:  
; APPLICANT: Steven P. Fling

## RESULT 8

US-09-598-419-176  
; Sequence 176, Application US/09598419  
; Patent No. 6565856  
; GENERAL INFORMATION:  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Scholler, John

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
; FILE REFERENCE: 210121.469C7  
; CURRENT APPLICATION NUMBER: US/09/620.412C  
; CURRENT FILING DATE: 2000-07-20  
; NUMBER OF SEQ ID NOS: 363  
; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
; SEQ ID NO 176  
; LENGTH: 982  
; TYPE: PRT  
; ORGANISM: Chlamydia  
; NAME/KEY: VARIANT  
; LOCATION: (1)---(982)  
; OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-620-412C-176

Query Match 4.2%; Score 122.5; DB 4; Length 982;  
Best Local Similarity 21.4%; Pred. No. 0.016;  
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;  
QY 36 PDQCSG--INRFVALFAVPLLSFHF--STNDPPAMNLRFLAADTLQKV----- 80  
Db 48 PLSCFNLGSGFTVLGRGHSILTFENIRTSNGAALS---SAADGLFTIEGKELSFNC 104  
QY 81 -AVLALL--ALASRGLSPRALGLDWSITLFSLSLTPNLTVMG-----IPLRGMYGA 130  
Db 105 NSLLAVLPAATTKGSGTPTTSTPSNGTTSYKTDL---LLNNEKFSFVSNLVSQDGA 161  
QY 131 SSAGTLMGVVVLQCIWYLMFLFEYRAARALVLDQPDGAAASIV-SRVDSDVVS 189  
Db 162 IDAKSLTVQGISKLCV-----FOENTAQA-----DGAQCVVTSFAMANEAPI 205  
QY 190 ARGDELEAEPPGVAGAGAVSSRGDAGRVRVTVRKSTSSRSB-AACSHSHSQTWQPRVS 248  
Db 206 A-----FVANVAGVRGGIAAVDQCGQ-----VSSSTSTEDPVVSFRNTAVEFDGNA 255  
QY 249 NLSGVEIYSLQ-----SSRNPTRGSSFNHADFENIVGAAAKG 286  
Db 256 RVGG-GIYSGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGQASNTSNYDGGGAI FCK 313  
QY 287 GGAAGDEBK-----ACGGGGHSPQQA VAVPAKRDHMLVWSSS 330  
Db 314 NGAQAGSNNSGVSFDGEGVVFSSNVAAGKG-----AIYAKKLSV-----AN 357  
QY 331 ASPV-----SRAAVHVFAG-----GADHADVLAKG-----AQAYDEYGRDDYSR 372  
Db 358 CGPVQFLNIANDGGAIIYLGESGELSADYGIIFDGNLKRKTAKENAADVNGVTVSQA 417  
QY 373 TNGSGG-----ADKGG-----PTLSKLSNSTAQ-----LYPKDDGEGRAAAVAMPAS 417  
Db 418 ISMSGGKITTIRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTDIVFANGS 477  
QY 418 VMTRLILIMWRKLI RNPNTYSSLGIVVMSLVSYRWGIEMPAIAR-----SISILSDAG 472  
Db 478 -----STLYQNVITIEQGRIVLRKAKLSVNSLSQTG 508  
QY 473 --LGMAMFSLGLFVALQPRIIACGNKLAIAIANGVRFVAGPAVMAAASIAVGLRGVLLHIA 530  
Db 509 GSLYMEAGSTLDFVTQP-----PQPPAANQLITLSN--LHLS 545  
QY 531 IVQAAALPQGI V-PFVFAKEYGVHPDIL-STAYGPTITSHG 567  
Db 546 LSSLLANNVNTNPTNPPAQDSDHPAVIGSTTAGSVTISG 584

; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
 ; FILE REFERENCE: 210121.469C6  
 ; CURRENT APPLICATION NUMBER: US/09/598,419  
 ; CURRENT FILING DATE: 2000-06-20  
 ; NUMBER OF SEQ ID NOS: 357  
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 ; SEQ ID NO 176  
 ; LENGTH: 982  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia  
 ; FEATURES:  
 ; NAME/KEY: VARIANT  
 ; LOCATION: (1)...(982)  
 ; OTHER INFORMATION: Xaa = Any Amino Acid  
 US-09-598-419-176

Query Match 4.2%; Score 122.5; DB 4; Length 982;  
 Best Local Similarity 21.4%; Pred. No. 0.016;  
 Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;  
 Qy 36 PDQCSG--INRFVALFAVPLLSFHF1--STNDPFAMNLFRLAADTLQKV----- 80  
 Db 48 PLSCFNLGSGFTVLGRGHSITFENIRTSNGAALSN---SAADGLFTIEGFKELSFNC 104  
 Qy 81 -AVLALL--ALASRGLSSPRALGLDWSITLPSLSTPLNTLVMG-----IPLLRGMYGA 130  
 Db 105 NSLLAVLPAAATTKNGSQPTTTTSPNGTIYSKTDL---LLNNEKFSFYSNLVSGDGA 161  
 Qy 131 SSAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQPDGAAASIV--SPRVSDVVSL 189  
 Db 162 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVVTFSAMANEAPI 205  
 Qy 190 ARGDVELEAPDGVAGAVSSRGDAGRVRVVRKSTSSRSE--AACSHSHSQTMQPRVS 248  
 Db 206 A-----FVANVAGVRGGIAAIVDQCGG-----VSSSTSTEDPVVFSRNTAVEFDGNA 255  
 Qy 249 NLGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAAAKG 286  
 Db 256 RVGG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT--SGQASNTSNNGYDGGAI FCK 313  
 Qy 287 GCGAAGDEBK-----ACGGGGGHSPOQAVAVPAKEDLHMLVWSSS 330  
 Db 314 NGAQAGNNSGVSFDGEGVVFSSNVAAGKG-----AIYAKKLSV-----AN 357  
 Qy 331 ASPV-----SERAHVHVGAG---GADHADVLAKG-----AQAYDEYGRDDYSSR 372  
 Db 358 CGPVQFLRNANDGGAIVLGESGELSADYGDIIFDGNLKRKTAKENAAADVNGVTSSQA 417  
 Qy 373 TKNGSGG-----ADKGG-----PTLSKLGSNSTAQ---LYPKDDGEGRAAAVAMPAS 417  
 Db 418 ISMGSGKITTIRAKAGHQILFNDPIEMANGNQPQSSKLLKINDGEGYTDIVFANGS 477  
 Qy 418 VMTRLILMVRKLIIRNPNTYSSLIGVWVSLVSYRWGIEMPAIAR-----SISILSDAG 472  
 Db 478 -----STLYQNVITIEQGRIVLREKAKLSVNSLQSG 508  
 Qy 473 --LGMAMFSLGLFMALQPRILACGNKLAIAIANGVRFPVAGPAAVMAAASIAVGLRGVLLHIA 530  
 Db 509 GSLYWEAGSLDFVTFQP-----POQPPAANQLITLSN--LHLS 545  
 Qy 531 IVQAALPQGIIV-PFVFAKEYGVHPDIL-STAYGPITSHG 567  
 Db 546 LSSLLANNAVTPPTNPAPQDSHPAVIGSTTAGSVTISG 584

## RESULT 9

US-09-556-877-190  
 ; Sequence 190, Application US/09556877  
 ; Patent No. 6432916  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Probst, Peter  
 ; APPLICANT: Bhatia, Ajay

## RESULT 10

US-09-620-412C-190  
 ; Sequence 190, Application US/09620412C  
 ; Patent No. 6448234  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Steven P. Fling  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION

; APPLICANT: Skeiky, Yasir  
 ; APPLICANT: Fling, Steve  
 ; APPLICANT: Maisonneuve, Jeff  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
 ; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
 ; FILE REFERENCE: 210121.469C5  
 ; CURRENT APPLICATION NUMBER: US/09/556,877  
 ; CURRENT FILING DATE: 2000-04-19  
 ; NUMBER OF SEQ ID NOS: 305  
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0  
 ; SEQ ID NO 190  
 ; LENGTH: 1006  
 ; TYPE: PRT  
 ; ORGANISM: Chlamydia  
 US-09-556-877-190

Query Match 4.2%; Score 122.5; DB 4; Length 1006;  
 Best Local Similarity 21.4%; Pred. No. 0.016;  
 Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;  
 Qy 36 PDQCSG--INRFVALFAVPLLSFHF1--STNDPFAMNLFRLAADTLQKV----- 80  
 Db 72 PLSCFNLGSGFTVLGRGHSITFENIRTSNGAALSN---SAADGLFTIEGFKELSFNC 128  
 Qy 81 -AVLALL--ALASRGLSSPRALGLDWSITLPSLSTPLNTLVMG-----IPLLRGMYGA 130  
 Db 129 NSLLAVLPAAATTKNGSQPTTTTSPNGTIYSKTDL---LLNNEKFSFYSNLVSGDGA 185  
 Qy 131 SSAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQPDGAAASIV--SPRVSDVVSL 189  
 Db 186 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVVTFSAMANEAPI 229  
 Qy 190 ARGDVELEAPDGVAGAVSSRGDAGRVRVVRKSTSSRSE--AACSHSHSQTMQPRVS 248  
 Db 230 A-----FVANVAGVRGGIAAIVDQCGG-----VSSSTSTEDPVVFSRNTAVEFDGNA 279  
 Qy 249 NLGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAAAKG 286  
 Db 280 RVGG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT--SGQASNTSNNGYDGGAI FCK 337  
 Qy 287 GCGAAGDEBK-----ACGGGGGHSPOQAVAVPAKEDLHMLVWSSS 330  
 Db 338 NGAQAGNNSGVSFDGEGVVFSSNVAAGKG-----AIYAKKLSV-----AN 381  
 Qy 331 ASPV-----SERAHVHVGAG---GADHADVLAKG-----AQAYDEYGRDDYSSR 372  
 Db 382 CGPVQFLRNANDGGAIVLGESGELSADYGDIIFDGNLKRKTAKENAAADVNGVTSSQA 441  
 Qy 373 TKNGSGG-----ADKGG-----PTLSKLGSNSTAQ---LYPKDDGEGRAAAVAMPAS 417  
 Db 442 ISMGSGKITTIRAKAGHQILFNDPIEMANGNQPQSSKLLKINDGEGYTDIVFANGS 501  
 Qy 418 VMTRLILMVRKLIIRNPNTYSSLIGVWVSLVSYRWGIEMPAIAR-----SISILSDAG 472  
 Db 502 -----STLYQNVITIEQGRIVLREKAKLSVNSLQSG 532  
 Qy 473 --LGMAMFSLGLFMALQPRILACGNKLAIAIANGVRFPVAGPAAVMAAASIAVGLRGVLLHIA 530  
 Db 533 GSLYWEAGSLDFVTFQP-----POQPPAANQLITLSN--LHLS 569  
 Qy 531 IVQAALPQGIIV-PFVFAKEYGVHPDIL-STAYGPITSHG 567  
 Db 570 LSSLLANNAVTPPTNPAPQDSHPAVIGSTTAGSVTISG 608

```

; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620.412C
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-190

Query Match
Best Local Similarity 4.2%; Score 122.5; DB 4; Length 1006;
Best Local Similarity 21.4%; Pred. No. 0.016;
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

QY 36 PDQCSG--INRFVAFVAPVLLSFHFI--STNDPPAMNLRFLAADTLQKV----- 80
DB 72 PLSCFENLGSFTVLGRGHSUTFFENITSTNGAALSN---SAADGLTFIEGFKELSFNSC 128
QY 81 -AVLALL--ALASRLSSPRALGLDWSITLFSLTPLNTLVMG-----IPLRGMYGA 130
DB 129 NSLLAVLPAATNKGSGQPTTTSTPSNGTIYSKTDL---LLLNNEKFSFVSNLVSGDGA 185
QY 131 SSAGTLMVQVVVLOCILWYTLMLFLFEYRAARALVLDQFPDGAASIV-SRVDSDVSVL 189
DB 186 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVVTFSAMANEAPI 229
QY 190 ARGDVELAEPPGVAGAVSSRGDAGRVRVTVRKSTSSRSE--AACSHSHSQTWPQPRVS 248
DB 230 A-----FVANVAGVRGGIAAVDQCGQG-----VSSSTSTEDPVVFSRNTAVEFDGNA 279
QY 249 NLSGVEIYSLO-----SSRNPTPRGSSFNHADPFNIVGAAAG 286
DB 280 RVGG--GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT--SGQASNTSNNGYDGGAI 337
QY 287 GCGAAGDEBK-----ACGGGGGGHSPQQA VAVPAKRDJHMLVWSSS 330
DB 338 NGAQAGSNNGSVSFDGEGVFFSSNVAAGKGG-----AIYAKLSV-----AN 381
QY 331 ASPV-----SERAHVHFGAG-----GADHADVLAKG-----AAQAYDEYGRDDYSR 372
DB 382 CGVQFLRNANDGAIYLGESGELSADYGIIDFNGLKRTAKENAADVNGVTVSSQA 441
QY 373 TNGSGG-----ADKGG-----PTLSKLGSNSTAQ-----LYPKDGEGRAAA VAMPAS 417
DB 442 ISMGSGGKITTRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTDIVFANGS 501
QY 418 VMTRLILIMWRKLIRNPNTYSSLIGVWSLVSVRWGIEMPAIIAR-----SISILSDAG 472
DB 502 -----STLYQNVTIEQGRIVLREKAKLSVNSLSQTG 532
QY 473 --LGMAMFSLGLFMAQPRIIACGNKLAIAIANGVRFVAGPAVMAAASIAVGLRGVLLHIA 530
DB 533 GSLYMEAGSTLDFVTPOP-----PQPPAANQLITLSN--LHLS 569
QY 531 IVQAALPQGI V-PFVFAKEYGVHPDIL-STAYGPIITSHG 567
DB 570 LSSLLANNNAVTPNPPPAQDSHPAVIGSTTAGSVITSG 608

RESULT 11
US-09-598-419-190
; Sequence 190, Application US/09598419
; Patent No. 6565856
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir A.W.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C6
; CURRENT APPLICATION NUMBER: US/09/598.419
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 357

; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 190
; LENGTH: 1006
; TYPE: PRT
; ORGANISM: Chlamydia
US-09-620-412C-190

Query Match
Best Local Similarity 4.2%; Score 122.5; DB 4; Length 1006;
Best Local Similarity 21.4%; Pred. No. 0.016;
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

QY 36 PDQCSG--INRFVAFVAPVLLSFHFI--STNDPPAMNLRFLAADTLQKV----- 80
DB 72 PLSCFENLGSFTVLGRGHSUTFFENITSTNGAALSN---SAADGLTFIEGFKELSFNSC 128
QY 81 -AVLALL--ALASRLSSPRALGLDWSITLFSLTPLNTLVMG-----IPLRGMYGA 130
DB 129 NSLLAVLPAATNKGSGQPTTTSTPSNGTIYSKTDL---LLLNNEKFSFVSNLVSGDGA 185
QY 131 SSAGTLMVQVVVLOCILWYTLMLFLFEYRAARALVLDQFPDGAASIV-SRVDSDVSVL 189
DB 186 IDAKSLTVQGISKLCV-----FOENTAQA-----DGGACQVVTFSAMANEAPI 229
QY 190 ARGDVELAEPPGVAGAVSSRGDAGRVRVTVRKSTSSRSE--AACSHSHSQTWPQPRVS 248
DB 230 A-----FVANVAGVRGGIAAVDQCGQG-----VSSSTSTEDPVVFSRNTAVEFDGNA 279
QY 249 NLSGVEIYSLO-----SSRNPTPRGSSFNHADPFNIVGAAAG 286
DB 280 RVGG--GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT--SGQASNTSNNGYDGGAI 337
QY 287 GCGAAGDEBK-----ACGGGGGGHSPQQA VAVPAKRDJHMLVWSSS 330
DB 338 NGAQAGSNNGSVSFDGEGVFFSSNVAAGKGG-----AIYAKLSV-----AN 381
QY 331 ASPV-----SERAHVHFGAG-----GADHADVLAKG-----AAQAYDEYGRDDYSR 372
DB 382 CGVQFLRNANDGAIYLGESGELSADYGIIDFNGLKRTAKENAADVNGVTVSSQA 441
QY 373 TNGSGG-----ADKGG-----PTLSKLGSNSTAQ-----LYPKDGEGRAAA VAMPAS 417
DB 442 ISMGSGGKITTRAKAGHQILFNDPIEMANGNNQPAQSSKLLKINDGEGYTDIVFANGS 501
QY 418 VMTRLILIMWRKLIRNPNTYSSLIGVWSLVSVRWGIEMPAIIAR-----SISILSDAG 472
DB 502 -----STLYQNVTIEQGRIVLREKAKLSVNSLSQTG 532
QY 473 --LGMAMFSLGLFMAQPRIIACGNKLAIAIANGVRFVAGPAVMAAASIAVGLRGVLLHIA 530
DB 533 GSLYMEAGSTLDFVTPOP-----PQPPAANQLITLSN--LHLS 569
QY 531 IVQAALPQGI V-PFVFAKEYGVHPDIL-STAYGPIITSHG 567
DB 570 LSSLLANNNAVTPNPPPAQDSHPAVIGSTTAGSVITSG 608

RESULT 12
US-09-612-402B-16
; Sequence 16, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; TITLE OF INVENTION: CHLAMYDIA PROTEIN, GENE SEQUENCE AND USES THEREOF
; FILE REFERENCE: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612.402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 16

```

```

; LENGTH: 1013
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-16

Query Match
Best Local Similarity 4.2%; Score 122.5; DB 4; Length 1013;
Matches 123; Conservative 66; Mismatches 199; Indels 191; Gaps 30;

QY 81 AVALL--ALASGLSPRALGLDMSITLFSLTPLNTLVMG-----IPLLRMYGAS 131
Db 136 SLIAVLPAATNNGSOTPTTSPNGTIYSKTDL---LLLNKEKFSYSLNLSVSGGGT 192
QY 132 SAGTLMVQVVVLCIITMYLMLFFRYAARALVLDQFPDGAASIV-SFRVDSVWSLA 190
Db 193 DAKSLTVQGISKLCV-----FQENTAQA-----DGGACQVTSFSAMANEAPIA 236
QY 191 RGDVELEAEPPDGVAGAVSSRGDAGRVRVTRKSTSSRSE-AACSHSHSQTMQPRVSN 249
Db 237 -----FTANVAGVRGGGIAAVQDQGG-----VSSSTSTEDPVVFSRNTAVEFDGNVAR 286
QY 250 LSGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAA-AKG 286
Db 287 VCG-GIYSYGNVAFNLNGKTLFLNNVSPVYIAAQPT-NGQASNTSDNYGGAIFCKN 344
QY 287 GCGAAGDEKG-----ACGGGGGCHSPQQAIVAPAKRDLHMLVWSSS 330
Db 345 GAGAAGSNNSGSVFGEVGVFFSSNVAAGKGG-----AIYAKLSV-----AN 388
QY 331 ASPV-----SERAHVHFGAG-----GADHADVLAKG-----AQAYDEYGRDDYSSR 372
Db 389 CGPVQFLGNANDGGAIVLGESGELSADYGDIIFDGNLKRKTAKENAADVNGVTVSQA 448
QY 373 TKNGSG-----ADKGG-----PTLSKLGNSTAQYLPK-----DDGGRAAAAMPAS 417
Db 449 ISMGSGCKITLRAKAGHQLFNDPIEMANGNNQPAQSSEPLKINDGEGYGDIVFANG- 507
QY 418 VMTLILIMWRKILRNPTYSLLIGVWVSLVSRWGIEMPAIIAR-----SISILSDAG 472
Db 508 -----NSTLYQNV-----IEQGRIVLRERAKLSVNSLSQTG 539
QY 473 --LGMAMFSLGLFVALQPLIACGNKLAALANGVFPVAGPAVMAAASIAVGRVLLHIA 530
Db 540 GSYLMEAGSLDVTPTQP-----POPPAANQLITLSN--LHLS 576
QY 531 IVQALPQGIIV-PFVFAKEYGVHFDIL-STAYGPIITSHG 567
Db 577 LSSILANNAVTPNTPPAQDSHPAVIGSTTAGPVITSG 615

RESULT 13
US-09-612-402B-17
; Sequence 17, Application US/09612402B
; Patent No. 6642023
; GENERAL INFORMATION:
; APPLICANT: Jackson, W. James
; APPLICANT: Pace, John
; TITLE OF INVENTION: Chlamydia Protein, Gene Sequence and Uses Thereof
; FILE REFERENCES: 7969-086-999
; CURRENT APPLICATION NUMBER: US/09/612,402B
; CURRENT FILING DATE: 2000-07-06
; PRIOR APPLICATION NUMBER: 08/942,596
; PRIOR FILING DATE: 1997-10-02
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 17
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Chlamydia sp.
US-09-612-402B-17

Query Match
4.2%; Score 121; DB 4; Length 505;

Best Local Similarity 22.0%; Pred. No. 0.016;
Matches 125; Conservative 42; Mismatches 152; Indels 248; Gaps 28;

QY 124 LRGMYGASSAGTLMVQVVVLCIITMYLMLFFRYAARALVLDQFP-----DG-- 172
Db 859 LAGMFRATEBQRLQARLV-----TAPAVQLGDVPLARLAPAVADGAG 900
QY 173 -----AAASIVSFRVDSVVSILAR-----GDVELEAEPDGVAG-----AGAVSSR 212
Db 901 GRDQPAALVWDGLPLGVVVLAEVVRVGVGAQVAGRLERAAVFFQHQHRRHVGLAAV 960
QY 213 GGDAGRVRVTRKSTSSRSEACSHSHSQTMQPRVSNLSGVETYSLSOSSRNPTRGSSFN 272
Db 961 VDEVGARPVVE-----ELLQDHVTEGHAQRGVGALLGVE-----PEVGELG 1001

Best Local Similarity 22.8%; Pred. No. 0.008;
Matches 107; Conservative 56; Mismatches 158; Indels 148; Gaps 25;

QY 36 PDCSG--INRFVALFAVELLSFHF--STNDPFFAMNLFRLAADTLQKV----- 80
Db 50 PLSCFNGLLGSTVLGRGHSITFEINRTSTNGAALSN---SAADGLFTIEGKELSPENC 106
QY 81 -AVALL--ALASRGLSSPRALGLDMSITLFSLSLTPNTLVMG-----IPLLRMYGAS 130
Db 107 NSLIALVLPARTNKGSOPTTTSTPSNGTIYSKTDL---LLLNKEKFSYSLNLSVSGDGA 163
QY 131 SAGTLMVQVVVLCIITMYLMLFFRYAARALVLDQFPDGAASIV-SFRVDSVWSL 189
Db 164 IDAKSLTVQGISKLCV-----FQENTAQA-----DGGACQVTSFSAMANEAPI 207
QY 190 ARGDVELEAEPPDGVAGAVSSRGDAGRVRVTRKSTSSRSE-AACSHSHSQTMQPRVS 248
Db 208 A-----FVANVAGVRGGGIAAVQDQGG-----VSSSTSTEDPVVFSRNTAVEFDGNVA 257
QY 249 NLGVEIYSLQ-----SSRNPTRGSSFNHADFFNIVGAAAGK 286
Db 258 RVGG-GIYSYGNVAFNLNGKTLFLNNVSPVYIAAQPT-SQASNTSNNYCDGGAIFCK 315
QY 287 GCGAAGDEKG-----ACGGGGGCHSPQQAIVAPAKRDLHMLVWSSS 330
Db 316 NGAQAGSNNSGSVFGEVGVFFSSNVAAGKGG-----AIYAKLSV-----AN 359
QY 331 ASPV-----SERAHVHFGAG-----GADHADVLAKG-----AQAYDEYGRDDYSSR 372
Db 360 CGPVQFLGNANDGGAIVLGESGELSADYGDIIFDGNLKRKTAKENAADVNGVTVSQA 419
QY 373 TKNGSG-----ADKGG-----PTLSKLGNSTAQYLPK-----LYPKDDGEG 406
Db 420 ISMGSGCKITLRAKAGHQLFNDPIEMANGNNQPAQSSEPLKINDGEG 468

RESULT 14
US-09-252-991A-20266
; Sequence 20266, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCES: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20266
; LENGTH: 1501
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20266

Query Match
4.2%; Score 121; DB 4; Length 1501;

Best Local Similarity 22.0%; Pred. No. 0.041;
Matches 125; Conservative 42; Mismatches 152; Indels 248; Gaps 28;

QY 124 LRGMYGASSAGTLMVQVVVLCIITMYLMLFFRYAARALVLDQFP-----DG-- 172
Db 859 LAGMFRATEBQRLQARLV-----TAPAVQLGDVPLARLAPAVADGAG 900
QY 173 -----AAASIVSFRVDSVVSILAR-----GDVELEAEPDGVAG-----AGAVSSR 212
Db 901 GRDQPAALVWDGLPLGVVVLAEVVRVGVGAQVAGRLERAAVFFQHQHRRHVGLAAV 960
QY 213 GGDAGRVRVTRKSTSSRSEACSHSHSQTMQPRVSNLSGVETYSLSOSSRNPTRGSSFN 272
Db 961 VDEVGARPVVE-----ELLQDHVTEGHAQRGVGALLGVE-----PEVGELG 1001

```

QY 273 HADFFNIVGAAKGGG-----GAAGDEKAGCGGGGGH-----SP 308  
Db 1002 H---LGVIRGDDGLGALVAHLGEBVRVGRPRLRHVGAFCDDVAGVVPVGRLRHVGLLAP 1058  
QY 309 QPQA-----VAVPAKRDHMLVWSSSASPVSERAAVHVFGAGG-ADHADVLAKGAQAYDE 363  
Db 1059 DLRAGRQVAVP-----VVEAQADPADQR---QVAGAGGVADH-----R 1094  
QY 364 YGRD-----DYSSRTKNGSGADKGGPTLSKLGNSSTAQLYPKDDGEGRAAAVAMPASVM 419  
Db 1095 HGRDREADHPVTV-GLGGVDG-----RGDDLGLVPACTDEAAQALA----- 1139  
QY 420 TRLLIMVWRKLRNPNTYSLGVWVSLVSYRWGTEMPAIARSISILSDAGLGNMFPS 479  
Db 1140 -----LVALRLG-----GVLDLDAGPGVHRVA 1160  
QY 480 LGLEFALQPRI-----IACGNKLAAL-AMGVRFVAGPAVMAAAAI-----AVGL 522  
Db 1161 V-----LRPRLAPELQRLAHLRVLQAVGAVDVPFGIAGPARAAARLVVGQVRTGARVVGL 1215  
QY 523 RG-----VLLHI-----AIQVQALPGIIVP-----FVF 545  
Db 1216 LGFFDQAVLHVDLPGAGAGVHPVGGAHDLVVLPAIPVGTTPGAALVAGLAIVGEFAL 1275  
QY 546 -----AKXYGVHPDILSTAYGPITS 565  
Db 1276 LLBELQAVEQVTHLDLITRSVPGVTT 1302

RESULT 15  
US-09-252-991A-26256  
; Sequence 26256, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252.991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 26256  
; LENGTH: 542  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-26256

Query Match 4.1%; Score 119.5; DB 4; Length 542;  
Best Local Similarity 23.3%; Pred. No. 0.012;  
Matches 101; Conservative 40; Mismatches 158; Indels 135; Gaps 20;

QY 192 GDVLEAEPDGVAGAGVSRGGDAGKRVTVTKSTSSRSEACSHSHSQTWQPRVSNLS 251  
Db 10 GGVDVGGRRPG-----GAHGWNG-----TARPAHRPAPSQAQAAAD-- 45  
QY 252 GVEIYSLQSSRNPTFRGSSFNHADFFNIVGAAKGGGAAGDEKAGCGGGG----- 304  
Db 46 -----PRSPRRRAGACRRGRVQAQADPRPGGAAGAARRRPAACAGCAGAGAGV 96  
QY 305 ---GHSFPQQA---VAVPAK-----RKDLHMLVWSSSASPVSFSE-----RAAVH- 341  
Db 97 RLGRGSRQPAAPAAALFGRLAGPVAYPAAACRVGWHARPFGAQAAPAPAMEDPARAGGHH 156  
QY 342 ---VFGAGGADHADVLAKGAQAYDEVGRDDYSRTKNG--SGGADKGGPTLSKLGNSSTA 396  
Db 157 LRALAPATAQARALRAGG-----LSARLQHSRLAAGAACGGP-----VGAGRLA 202  
QY 397 QLYPKDDGEGRAAAVAMP-----PASVMTRLILIMVWRKLRNPNT 437

Db 203 R-----PRTVAVAMPARAGGPARAGLDGTLRTGACRTGPVDIAAGTCRALRPPV 254  
QY 438 YSSLIGVWVHSLVSYRWGTEMPAIARSISILSDA-----GLGNMFSLGLF 483  
Db 255 FRSAFS--QGSPLRLWALSPREAHMNFSELIOAVRRDPSSVVVPASWQGGRTATFG-GLV 311  
QY 484 MAL-----QPRIIACGNKLAALAMGVRFVAGPAVMAAAAIAGVLRGVLLHIAIVQALPQG 539  
Db 312 VALAYEAMLAIVVEAGRPLRSI--GVSVFGPLAPEQOPASFSARLLREGKAVSQVQVEVRQG 369  
QY 540 --IVPFFVFAKEYGV 551  
Db 370 EQVTVLQQA-SFGV 382

Search completed: March 3, 2004, 09:09:47  
Job time : 26 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 3, 2004, 09:08:49 ; Search time 53 Seconds  
(without alignments)  
2282.846 Million cell updates/sec

Title: US-10-030-884-14

Perfect score: 2910  
Sequence: 1 MITALDLXHLVAVPLVYA.....DLSTAYGPITSHGFITCS 573

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCT05\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129.5	4.5	1402	9	US-09-712-363-166
2	129	4.4	646	9	US-09-964-899-53
3	126	4.3	277	15	US-10-032-037B-25
4	126	4.3	277	15	US-10-029-988B-25
5	126	4.3	277	15	US-10-032-423A-25
6	125	4.3	462	10	US-09-919-039-324
7	124.5	4.3	623	13	US-10-108-605-125
8	124.5	4.3	623	13	US-10-108-605-129
9	122.5	4.2	982	9	US-09-841-132-176
10	122.5	4.2	1006	9	US-09-841-132-190
11	122	4.2	277	15	US-10-032-037B-203
12	122	4.2	277	15	US-10-029-988B-203
13	122	4.2	277	15	US-10-032-423A-203
14	122	4.2	323	15	US-10-374-780A-1400
15	121.5	4.2	585	9	US-09-738-626-6032
					Sequence 166, App
					Sequence 53, Appl
					Sequence 25, Appl
					Sequence 25, Appl
					Sequence 25, Appl
					Sequence 324, App
					Sequence 125, App
					Sequence 129, App
					Sequence 176, App
					Sequence 190, App
					Sequence 203, App
					Sequence 203, App
					Sequence 203, App
					Sequence 1400, Ap
					Sequence 6032, Ap

16	118	4.1	1176	14	US-10-156-761-13466	Sequence 13466, A
17	117	4.0	249	10	US-09-880-748-749	Sequence 749, App
18	116	4.0	249	10	US-09-880-748-2033	Sequence 2033, Ap
19	116	4.0	251	10	US-09-880-748-1047	Sequence 1047, Ap
20	115.5	4.0	246	10	US-09-880-748-928	Sequence 928, App
21	115	4.0	249	10	US-09-880-748-772	Sequence 772, App
22	115	4.0	249	10	US-09-880-748-833	Sequence 833, App
23	115	4.0	251	10	US-09-880-748-1053	Sequence 1053, Ap
24	115	4.0	251	10	US-09-880-748-1805	Sequence 1805, Ap
25	115	4.0	266	15	US-10-032-037B-204	Sequence 204, App
26	115	4.0	266	15	US-10-029-988B-204	Sequence 204, App
27	115	4.0	266	15	US-10-032-423A-204	Sequence 204, App
28	115	4.0	892	9	US-09-887-828A-2	Sequence 2, Appl
29	114.5	3.9	248	10	US-09-880-748-1643	Sequence 1643, Ap
30	114.5	3.9	248	10	US-09-880-748-1681	Sequence 1681, Ap
31	114.5	3.9	423	14	US-10-100-679-95	Sequence 95, Appl
32	114.5	3.9	471	14	US-10-100-679-96	Sequence 96, Appl
33	114.5	3.9	541	14	US-10-100-679-97	Sequence 97, Appl
34	114.5	3.9	582	14	US-10-100-679-81	Sequence 81, Appl
35	114.5	3.9	1482	10	US-09-940-316B-21	Sequence 21, Appl
36	114.5	3.9	1488	10	US-09-940-316B-17	Sequence 17, Appl
37	114	3.9	249	10	US-09-880-748-629	Sequence 629, App
38	114	3.9	251	10	US-09-880-748-1060	Sequence 1060, Ap
39	114	3.9	4471	14	US-10-205-032-10	Sequence 10, Appl
40	113.5	3.9	246	10	US-09-880-748-2068	Sequence 2068, Ap
41	113.5	3.9	248	10	US-09-880-748-1178	Sequence 1178, Ap
42	113.5	3.9	248	10	US-09-880-748-1306	Sequence 1306, Ap
43	113.5	3.9	248	10	US-09-880-748-1785	Sequence 1785, Ap
44	113.5	3.9	249	10	US-09-880-748-632	Sequence 632, App
45	113.5	3.9	249	10	US-09-880-748-673	Sequence 673, App

#### ALIGNMENTS

#### RESULT 1

US-09-712-363-166  
; Sequence 166, Application US/09712363  
; Patent No. US20020164588A1  
; GENERAL INFORMATION:  
; APPLICANT: Eisenberg, David  
; APPLICANT: Rotstein, Sergio H.  
; APPLICANT: Marcotte, Edward M.  
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
; FILE REFERENCE: 07419-032001  
; CURRENT APPLICATION NUMBER: US/09/712.363  
; CURRENT FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: PCT/US00/02246  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,531  
; PRIOR FILING DATE: 2000-02-01  
; PRIOR APPLICATION NUMBER: 60/117,844  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: 60/118,206  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: 60/126,593  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 60/134,093  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/134,092  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/165,124  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/165,086  
; PRIOR FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 292  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 166  
; LENGTH: 1402  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
US-09-712-363-166

Query Match 4.5%; Score 129.5; DB 9; Length 1402;  
Best Local Similarity 22.6%; Pred. No. 0.031;  
Matches 111; Conservative 53; Mismatches 189; Indels 139; Gaps 21;

QY 107 LPSLSTLPTNLVWGIPLLGKMGASAGTLMVQVVVLQCIWIYTLMLFEYRAARALVL 166  
DB 697 LFAVEVALYRLLMWGVRPGLVHSGVLAHAHAGALCLDPAAMLV-----AARGRLM 751  
QY 167 DQPPDGAASIVSFRVDSVWLSIARGDVELEA--EPDGVAGAGAVSSRGDAGRVRVTVR 224  
DB 752 QALPAGGAMFAVQAREE-EVAPWLGHDVSIAAVNGPASVVISGAHDAVSADRLRGQR 810  
QY 225 KSTSSRSEAAACHS-HSQTMQPRVSNLSGVE-----IYSSQSSRNPTPR--GSSFNEAD 275  
DB 811 RV-----HRLAVSHAFHSAHMEPIAFBFTAAAEALSGLTIPVINSVTGQLVADDFASAD 866  
QY 276 FF-NIVGAAAGGGGAAGDEEKA-----CGGGGGGHS-----PQOAVAVPAKRD-- 321  
DB 867 YNARHIRAVVRGDSVRSACAGASRFIEVGGGLTSLIEASLADAQIVSVPTLRKDRP 926  
QY 322 -----LHMLVSSASPVSEAAVH 341  
DB 927 EPVSVMTAAAGQFVSGMGLDWSVFSGYRPRKRVLPYAFQHKFWLAPAPSVSDPTAAG 986  
QY 342 VTGA--GGAD-----HAD-VLAKGAQAYDEV-----GRDDYSSTKNGSGG 379  
DB 987 QIGASGGAEELASGFAARLAGRSADQIAAEVVECHAAAVLGRDG-AGLDAGQAF 1045  
QY 380 ADKGGPTLS--KLGSNSTA-----QLYPKDDGEGRAAAVAMPP 415  
DB 1046 ADGGFNSLGAVALERNLTAVTATLPTATAIFDHPPTTELAQVLIQTIDGHGSSAAAAANP 1105  
QY 416 A-----SVMTRLILMV-----WRKLRLNPNTYSSLIGVMSVLSYRWGI-----BMP 458  
DB 1106 AERIDALTDLFQACDAGRDADGKXWVALASNTREMSFVRNNVSKVALLADGISDVV 1165  
QY 459 AIIARSISILSD 470  
DB 1166 VICIPLTVLSD 1177

RESULT 2  
US-09-964-899-53  
; Sequence 53, Application US/09964899  
; Patent No. US20020174446A1  
; GENERAL INFORMATION:  
; APPLICANT: Cohen, Dalia et al.  
; TITLE OF INVENTION: Identification of Genes Involved in  
; FILE REFERENCE: 4-31612 A  
; CURRENT APPLICATION NUMBER: US/09/964,899  
; CURRENT FILING DATE: 2001-09-27  
; PRIOR FILING DATE: 2001-09-29  
; PRIOR FILING DATE: 2000-09-29  
; PRIOR FILING DATE: 2001-06-14  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 53  
; LENGTH: 646  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-964-899-53

Query Match 4.4%; Score 129; DB 9; Length 646;  
Best Local Similarity 20.5%; Pred. No. 0.011;  
Matches 108; Conservative 43; Mismatches 153; Indels 224; Gaps 21;

QY 204 AGAGVSSRG-----DAGRVRVTVKSTSSRSEAAACHS-SHSQTMQPR 246  
DB 41 SGPGSGSPAGSTKPFVHAVPPSDPLRQANRLPIKVLKMLTART-----GHILHPEYLQPL 96

QY 247 VSN-LSCVEIYSLOSS-----RNPTPRGSSFNHADFFNIYVCAAKAGGGGAAGD- 293  
DB 97 PSTVPSPIELDAAKSPALALQAOTCSQTKGPDPPSPSS-KLSSVASNCGGAGGAGGAGDK 155  
QY 294 -----EEK-----GACGGGGGGHSPQOAVAVP 316  
DB 156 DTKSGPLKLSDIGVEDKSSPKYKPGSKKEPGGGGGGGGGGGGGGVSSEKSGFRVP 215  
QY 317 AKRKDLHMLVWSSASPVSEAAVHVFGAGADHADVLAKGAQAYDEYGRDDYSSRTKNG 376  
DB 216 -----SATCOPFTPTGTSPSSSASACSPGMLSLSAGGAPE--GKDDKXKDTDVG 262  
QY 377 SG-----GADKGGPTLSKLGSN 393  
DB 263 GKGKTGASREGGPTGLAGRI SCGGGINVDVNHQPDGPGGKXALGSDCGSSGSSSGSG 322  
QY 394 STAQLYKPDGEGRAAAVA-----MPPASVMTRLILIMVWRKLRNPNTY-SSLI 442  
DB 323 PSAPTSSSVLGSGLVAPVSPYKPGQTVPFLPPAGM-----TYPGSLA 364  
QY 443 GVV-----WSLSYRWGIEMPAIIARSISILSDAGLGMAMFSLGLF 483  
DB 365 GAYAGYPPQFLPHGVLDPTKPSLV---GAQLAAAAGSLGCSKPAG-----SSPLA 414  
QY 484 MALQPRIIA-----CONKLAATAMGVRFVAGPAMVMAASIAVG----- 521  
DB 415 GASPPSVMTASLCKRDPYCLSYHCASHLAGAAAASACAHDPAAAAAALKSGYPLVYPTHP 474  
QY 522 LRGVLLHIAIVQALPGIIVFVFAKEYGVHPDILSTAYGPIITSHGPI 569  
DB 475 LHGV--HSSLTAAAAAGATPP-----SLAGHPLYPYGP 506

RESULT 3  
US-10-032-037B-25  
; Sequence 25, Application US/10032037B  
; Publication No. US20040001822A1  
; GENERAL INFORMATION:  
; APPLICANT: Bio-Technology General Corp.  
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED  
; FILE REFERENCE: 10793/44  
; CURRENT APPLICATION NUMBER: US/10/032,037B  
; CURRENT FILING DATE: 2001-12-31  
; PRIOR FILING DATE: 2000-12-29  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 25  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-032-037B-25

Query Match 4.3%; Score 126; DB 15; Length 277;  
Best Local Similarity 22.3%; Pred. No. 0.0063;  
Matches 68; Conservative 36; Mismatches 125; Indels 76; Gaps 9;

QY 152 MLFLFEYRAARALVLDQFPDGAASIV-----SERVSDVVSILAR 191  
DB 1 MKYLLPTAAAGLLLLAAQPAWAEVOLVESGGVVRPGSLRLSCAASGFTDDYGMXSVR 60  
QY 192 GDVELEAEPDGVAGAGAVSSRGDAGRVRVTVRKSTSSRSEAAACHSHSQTMQPRVSNLS 251  
DB 61 -----QAPGKLEWVSGINWNGGSGTGYADSVKGRFTTISRDNA-----KNSLYLQWNSLRAE 111  
QY 252 GVEIYSIQSSRNPTPRGSSFNHADFFNIYVCAAKAGGGAGGAGDEKAGCGGGGGHSPQP 311  
DB 112 DTAVTYCARMAPVWQ-----GTLVTVSRGGGGSGG-----GGGGGGSSSELTQDP 159  
QY 312 AVAV-----PAKRLHMLVWSSASPVSEAAVHVFGA 345  
DB 160 AVSVALGQTVITCGDLSRSYASVYQKPGQAPVLVIYKNNRPSGIDRFS-----GS 215

```

QY 346 GGADHADVLAKGAQAVDEYGRDDY--SSRTKNGSGADKGGFTLSKLSNSTAQLYPKDD 403
Db 216 SSGNTASLTITGAQAED---ADYYCNSRDSGNHVFVGGGTKLTVLGAAGAEQKLISEED 272
QY 404 GEGRA 408
Db 273 LINGAA 277

```

## RESULT 4

```

US-10-029-988B-25
; Sequence 25, Application US/10029988B
; Publication No. US20040001839A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE OF INVENTION: MOLETTIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/45
; CURRENT APPLICATION NUMBER: US/10/029,988B
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-988B-25

```

```

Query Match 4.3%; Score 126; DB 15; Length 277;
Best Local Similarity 22.3%; Pred. No. 0.0063;
Matches 68; Conservative 36; Mismatches 125; Indels 76; Gaps 9;

```

```

QY 152 MLFLFEYRAARALVLDQFPDGAASIV-----SFRVDSVWVSLAR 191
Db 1 MKYLLPTAAAGLLLLAAQAPAMAEVOLVESGGVVRPGGSLRLSCAASGFTDDYGMWVR 60
QY 192 GDVELEAEPDGVAGAGAVSSRGGDAGRVRVTRKSTSRSEACSHSHSQTMQPRVSNLS 251
Db 61 -----QAPKGLEWVSGINWNGGSTGYADSVKGRFTISRDN-----KNSLYLQNSLRAE 111
QY 252 GVEIYLSQSSRNTPRGSSFNHADFFNIVGAAAGGGGAAGDEKAGCGGGGGHSPQPQ 311
Db 112 DTAVVYICARMRAPVIWQ-----GTLVTVSRGGGGSGG-----GGSGGGSSSELTQDP 159
QY 312 AVAV-----PAKRXDLHMLVWSSASPVSEAAVHVFGA 345
Db 160 AVSVALGQTVRITCQDLSRSYASWYQKPGQAPVLIYKNNRPSGIPDRFS-----GS 215
QY 346 GGADHADVLAKGAQAVDEYGRDDY--SSRTKNGSGADKGGFTLSKLSNSTAQLYPKDD 403
Db 216 SSGNTASLTITGAQAED---ADYYCNSRDSGNHVFVGGGTKLTVLGAAGAEQKLISEED 272
QY 404 GEGRA 408
Db 273 LINGAA 277

```

## RESULT 5

```

US-10-032-423A-25
; Sequence 25, Application US/10032423A
; Publication No. US20040002450A1
; GENERAL INFORMATION:
; APPLICANT: Bio-Technology General Corp.
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED
; FILE OF INVENTION: MOLETTIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF
; FILE REFERENCE: 10793/45
; CURRENT APPLICATION NUMBER: US/10/032,423A
; CURRENT FILING DATE: 2001-12-31
; PRIOR APPLICATION NUMBER: 60/258,948
; PRIOR FILING DATE: 12/29/2000

```

```

; NUMBER OF SEQ ID NOS: 204
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-032-423A-25

```

```

Query Match 4.3%; Score 126; DB 15; Length 277;
Best Local Similarity 22.3%; Pred. No. 0.0063;
Matches 68; Conservative 36; Mismatches 125; Indels 76; Gaps 9;
QY 152 MLFLFEYRAARALVLDQFPDGAASIV-----SFRVDSVWVSLAR 191
Db 1 MKYLLPTAAAGLLLLAAQAPAMAEVOLVESGGVVRPGGSLRLSCAASGFTDDYGMWVR 60
QY 192 GDVELEAEPDGVAGAGAVSSRGGDAGRVRVTRKSTSRSEACSHSHSQTMQPRVSNLS 251
Db 61 -----QAPKGLEWVSGINWNGGSTGYADSVKGRFTISRDN-----KNSLYLQNSLRAE 111
QY 252 GVEIYLSQSSRNTPRGSSFNHADFFNIVGAAAGGGGAAGDEKAGCGGGGGHSPQPQ 311
Db 112 DTAVVYICARMRAPVIWQ-----GTLVTVSRGGGGSGG-----GGSGGGSSSELTQDP 159
QY 312 AVAV-----PAKRXDLHMLVWSSASPVSEAAVHVFGA 345
Db 160 AVSVALGQTVRITCQDLSRSYASWYQKPGQAPVLIYKNNRPSGIPDRFS-----GS 215
QY 346 GGADHADVLAKGAQAVDEYGRDDY--SSRTKNGSGADKGGFTLSKLSNSTAQLYPKDD 403
Db 216 SSGNTASLTITGAQAED---ADYYCNSRDSGNHVFVGGGTKLTVLGAAGAEQKLISEED 272
QY 404 GEGRA 408
Db 273 LINGAA 277

```

## RESULT 6

```

US-09-919-039-324
; Sequence 324, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 324
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 1813444CD1
US-09-919-039-324

```

```

Query Match 4.3%; Score 125; DB 10; Length 462;
Best Local Similarity 27.1%; Pred. No. 0.016;
Matches 56; Conservative 18; Mismatches 75; Indels 58; Gaps 10;
QY 208 AVSSRGGDAGRVRVTRKSTSRSEACSHSHSQTMQPRVSNLSGVRIYSLQSRNTPR 267
Db 105 APSSTSGSYG-----SSQSSSYGQPSGYSQPSYGGQ--QSYGQQSYNP-PQ 153
QY 268 GGSFNHADFFNIVGAAAGGGGAAGDEKAGCGGGGGHSPQPQAVAPAKRDLMLVW 327
Db 154 G--YGGQNYN-----SSSGGGG-----GGGGGNYGQQ----- 181
QY 328 SSSASPVSEAAVHVFGAGGADHADVLAKGAQAVDEYGRDDYSSRTKNGSGADKGGPTL 387

```



NAME/KEY: VARIANT  
LOCATION: (1)...(982)  
OTHER INFORMATION: Xaa = Any Amino Acid  
US-09-841-132-176

Query Match  
Best Local Similarity 4.2%; Score 122.5; DB 9; Length 982;  
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

QY 36 PDQCSG--INRFVAFVPLLSFHF--STNDPFFAMNLFRLAADTLQKV-----80  
DB 48 PLSCFGNLLGSFTVLGRGHSITFNIRTSNGAALN---SAADGLFTIEGFKELSFNSC 104

QY 81 -AVALL--ALASRGLSSPRALGLDWSITLFSLSSTLPTNLVWG-----1PLLSGMVGA 130  
DB 105 NSLLAVLPAATNKGSTPTTSTPSNGTIYSKTDL---LLNNEKFSFYSNLVSGDGA 161

QY 131 SSAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQPDGAAASIV-SFRVSDVVS 189  
DB 162 IDAKSLTVQIGSKLCV-----FOENTAQA-----DGGACQVTSFSAMANEAPI 205

QY 190 ARGVLELEAPDGVAGAVSSRGDAGRVRVTRKSTSRSE-AACSHSHSQTWQPRVS 248  
DB 206 A-----FVANVAGVRGGGIAAQQDQGG-----VSSSTSTEDPVVSFRNTAVEFDGVA 255

QY 249 NLGSGVEIYSLQ-----SSRNPPTGRSSFNHADFFENIVGAAAKG 286  
DB 256 RVGG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGAASNTSNNGYDGGAI FCK 313

QY 287 GGGAGDEEKG-----ACGGGGGHSPOQAVAVPAKEKDLHMLVWSSS 330  
DB 314 NGAAGSNNSGVSFDGEGVVFSSNVAAGKG-----AIYAKKLSV-----AN 357

QY 331 ASPV-----SRAAVHVFAG-----GADHADVLAKG-----AAQVDEYGRDDYSR 372  
DB 358 CGPQVFLRNTANDGAIYLGESGELSADYGDIIIPDGNLKRKTAKENADVNGVTSSQA 417

QY 373 TKNSGG-----ADKGG-----PTLSKLGSNSTAQ-----LYPKDDEGRAAAVAMPAS 417  
DB 418 ISMGSGGKITTLRAKAGHQILFNDPIEMANGNQPACSSKLLKINDGEGYTGDI VFANGS 477

QY 418 VMTRLILIMVWRKLIERNPTYSSLIGVWVSLVSRWGIEMPAIAR-----SISILSDAG 472  
DB 478 -----STIQNTVIEQGRIVLRREKAKLSVNSLSQTG 508

QY 473 --LGMAMFSLGLFMALQPRIIACGNKLAIAIANGVRFVAGPAAVMAAASIAVGLRGVLLHIA 530  
DB 509 GSLYMEAGSTLDFVTQP-----PQOPPAANQLITLSN--LHLS 545

QY 531 IVQAALPQGIIV-PFVFAKEYGVHPDIL-STAYGPITSHG 567  
DB 546 LSSLLANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISG 584

RESULT 10  
US-09-841-132-190  
Sequence 190, Application US/09841132  
Patent No. US20020061848A1  
GENERAL INFORMATION:  
APPLICANT: Bhacia, Ajay  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Probst, Peter  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND  
TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION  
FILE REFERENCE: 210121.469C8  
CURRENT APPLICATION NUMBER: US/09/841.132  
CURRENT FILING DATE: 2001-04-23  
NUMBER OF SEQ ID NOS: 599  
SOFTWARE: FastSeq for Windows Version 3.0/4.0  
SEQ ID NO 190  
LENGTH: 1006  
TYPE: PRT  
ORGANISM: Chlamydia

US-09-841-132-190

Query Match  
Best Local Similarity 4.2%; Score 122.5; DB 9; Length 1006;  
Matches 137; Conservative 71; Mismatches 222; Indels 209; Gaps 32;

QY 36 PDQCSG--INRFVAFVPLLSFHF--STNDPFFAMNLFRLAADTLQKV-----80  
DB 72 PLSCFGNLLGSFTVLGRGHSITFNIRTSNGAALN---SAADGLFTIEGFKELSFNSC 128

QY 81 -AVALL--ALASRGLSSPRALGLDWSITLFSLSSTLPTNLVWG-----1PLLSGMVGA 130  
DB 129 NSLLAVLPAATNKGSTPTTSTPSNGTIYSKTDL---LLNNEKFSFYSNLVSGDGA 185

QY 131 SSAGTLMVQVVLQCIWYTLMLFLFEYRAARALVLDQPDGAAASIV-SFRVSDVVS 189  
DB 186 IDAKSLTVQIGSKLCV-----FOENTAQA-----DGGACQVTSFSAMANEAPI 229

QY 190 ARGVLELEAPDGVAGAVSSRGDAGRVRVTRKSTSRSE-AACSHSHSQTWQPRVS 248  
DB 230 A-----FVANVAGVRGGGIAAQQDQGG-----VSSSTSTEDPVVSFRNTAVEFDGVA 279

QY 249 NLGSGVEIYSLQ-----SSRNPPTGRSSFNHADFFENIVGAAAKG 286  
DB 280 RVGG-GIYSYGNVAFNLNGKTLFLNNVASPVYIAAKQPT-SGAASNTSNNGYDGGAI FCK 337

QY 287 GGGAGDEEKG-----ACGGGGGHSPOQAVAVPAKEKDLHMLVWSSS 330  
DB 338 NGAAGSNNSGVSFDGEGVVFSSNVAAGKG-----AIYAKKLSV-----AN 381

QY 331 ASPV-----SRAAVHVFAG-----GADHADVLAKG-----AAQVDEYGRDDYSR 372  
DB 382 CGPQVFLRNTANDGAIYLGESGELSADYGDIIIPDGNLKRKTAKENADVNGVTSSQA 441

QY 373 TKNSGG-----ADKGG-----PTLSKLGSNSTAQ-----LYPKDDEGRAAAVAMPAS 417  
DB 442 ISMGSGGKITTLRAKAGHQILFNDPIEMANGNQPACSSKLLKINDGEGYTGDI VFANGS 501

QY 418 VMTRLILIMVWRKLIERNPTYSSLIGVWVSLVSRWGIEMPAIAR-----SISILSDAG 472  
DB 502 -----STIQNTVIEQGRIVLRREKAKLSVNSLSQTG 532

QY 473 --LGMAMFSLGLFMALQPRIIACGNKLAIAIANGVRFVAGPAAVMAAASIAVGLRGVLLHIA 530  
DB 533 GSLYMEAGSTLDFVTQP-----PQOPPAANQLITLSN--LHLS 569

QY 531 IVQAALPQGIIV-PFVFAKEYGVHPDIL-STAYGPITSHG 567  
DB 570 LSSLLANNVNTNPTNPPAQDSHPAVIGSTTAGSVTISG 608

RESULT 11  
US-10-032-037B-203  
Sequence 203, Application US/10032037B  
Publication No. US20040001822A1  
GENERAL INFORMATION:  
APPLICANT: Bio-Technology General Corp.  
TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED  
TITLE OF INVENTION: MOETIES, ANTIBODIES TO SUCH EPITOPES, AND USES THEREOF  
FILE REFERENCE: 10793/44  
CURRENT APPLICATION NUMBER: US/10/032.037B  
CURRENT FILING DATE: 2001-12-31  
PRIOR APPLICATION NUMBER: 60/258,948  
PRIOR FILING DATE: 2000-12-29  
NUMBER OF SEQ ID NOS: 204  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 203  
LENGTH: 277  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-032-037B-203

Query Match  
4.2%; Score 122; DB 15; Length 277;

Best Local Similarity 21.8%; Pred. No. 0.014;  
Matches 67; Conservative 40; Mismatches 120; Indels 80; Gaps 9;  
QY 152 MLFLFEYRAARALVLDQFPDGAASIV-----SPRVSDVVSLAR 191  
Db 1 MKYLLPTAAAGLLLLAAQPAMAEVQLVESGGGVVRRPGGSLRLSCAASGFTFDDYGNMSVVR 60  
QY 192 GDVELEAEAPDGAGAGAVSSRGGDAGRVVTRKSTSSRSEAAACSHSHSQTMQPRVSNLS 251  
Db 61 -----QAPCKGLEWVSGINWNGSGTGADSVKGRFTISRDN-----KNSL 101  
QY 252 GVEIYSLQSSRNPTPRGSSFNHADFF--NIVGAAKGGGGAAGDEBKACGGGGHSPQ 309  
Db 102 YLQMNLSRAEDTAVYICARLTHPYFWGQGLTVTVSRGGGGSGG---GGSGGGGSSSELTQ 157  
QY 310 PQAVAV-----PAKRKDLHMLVWSSSASPVSERAHVHF 343  
Db 158 DPAVSVALGQTVITTCQDLSRYSYASWYQKPGQAPVLIYKNNRPSGIPDRFS----213  
QY 344 GAGGADHADVLAKGAQAYDEYGRDYY--SSRTKNGSGGADKGGFTLSKLGNSSTAQLYPK 401  
Db 214 GSSSGNTASLTITGAQAED-----ADYYCNSRDSGNHVVFGGQTKLTVLGAAGAEQKLISE 270  
QY 402 DDGEGR 408  
Db 271 EDLNGAA 277

RESULT 12  
US-10-029-988B-203  
; Sequence 203, Application US/10029988B  
; Publication No. US20040001839A1  
; GENERAL INFORMATION:  
; APPLICANT: Bio-Technology General Corp.  
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED  
; FILE REFERENCE: 10793/46  
; CURRENT APPLICATION NUMBER: US/10/029,988B  
; PRIOR FILING DATE: 2001-12-31  
; PRIOR FILING DATE: 2000-12-29  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 203  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-029-988B-203

Query Match 4.2%; Score 122; DB 15; Length 277;  
Best Local Similarity 21.8%; Pred. No. 0.014;  
Matches 67; Conservative 40; Mismatches 120; Indels 80; Gaps 9;  
QY 152 MLFLFEYRAARALVLDQFPDGAASIV-----SPRVSDVVSLAR 191  
Db 1 MKYLLPTAAAGLLLLAAQPAMAEVQLVESGGGVVRRPGGSLRLSCAASGFTFDDYGNMSVVR 60  
QY 192 GDVELEAEAPDGAGAGAVSSRGGDAGRVVTRKSTSSRSEAAACSHSHSQTMQPRVSNLS 251  
Db 61 -----QAPCKGLEWVSGINWNGSGTGADSVKGRFTISRDN-----KNSL 101  
QY 252 GVEIYSLQSSRNPTPRGSSFNHADFF--NIVGAAKGGGGAAGDEBKACGGGGHSPQ 309  
Db 102 YLQMNLSRAEDTAVYICARLTHPYFWGQGLTVTVSRGGGGSGG---GGSGGGGSSSELTQ 157  
QY 310 PQAVAV-----PAKRKDLHMLVWSSSASPVSERAHVHF 343  
Db 158 DPAVSVALGQTVITTCQDLSRYSYASWYQKPGQAPVLIYKNNRPSGIPDRFS----213  
QY 344 GAGGADHADVLAKGAQAYDEYGRDYY--SSRTKNGSGGADKGGFTLSKLGNSSTAQLYPK 401  
Db 214 GSSSGNTASLTITGAQAED-----ADYYCNSRDSGNHVVFGGQTKLTVLGAAGAEQKLISE 270

QY 402 DDGEGR 408  
Db 271 EDLNGAA 277  
RESULT 13  
US-10-032-423A-203  
; Sequence 203, Application US/10032423A  
; Publication No. US20040002450A1  
; GENERAL INFORMATION:  
; APPLICANT: Bio-Technology General Corp.  
; TITLE OF INVENTION: Y17-ISOLATED MOLECULES COMPRISING EPITOPES CONTAINING SULFATED  
; FILE REFERENCE: 10793/45  
; CURRENT APPLICATION NUMBER: US/10/032,423A  
; PRIOR FILING DATE: 2001-12-31  
; PRIOR FILING DATE: 2000-12-29  
; NUMBER OF SEQ ID NOS: 204  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 203  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-032-423A-203

Query Match 4.2%; Score 122; DB 15; Length 277;  
Best Local Similarity 21.8%; Pred. No. 0.014;  
Matches 67; Conservative 40; Mismatches 120; Indels 80; Gaps 9;  
QY 152 MLFLFEYRAARALVLDQFPDGAASIV-----SPRVSDVVSLAR 191  
Db 1 MKYLLPTAAAGLLLLAAQPAMAEVQLVESGGGVVRRPGGSLRLSCAASGFTFDDYGNMSVVR 60  
QY 192 GDVELEAEAPDGAGAGAVSSRGGDAGRVVTRKSTSSRSEAAACSHSHSQTMQPRVSNLS 251  
Db 61 -----QAPCKGLEWVSGINWNGSGTGADSVKGRFTISRDN-----KNSL 101  
QY 252 GVEIYSLQSSRNPTPRGSSFNHADFF--NIVGAAKGGGGAAGDEBKACGGGGHSPQ 309  
Db 102 YLQMNLSRAEDTAVYICARLTHPYFWGQGLTVTVSRGGGGSGG---GGSGGGGSSSELTQ 157  
QY 310 PQAVAV-----PAKRKDLHMLVWSSSASPVSERAHVHF 343  
Db 158 DPAVSVALGQTVITTCQDLSRYSYASWYQKPGQAPVLIYKNNRPSGIPDRFS----213  
QY 344 GAGGADHADVLAKGAQAYDEYGRDYY--SSRTKNGSGGADKGGFTLSKLGNSSTAQLYPK 401  
Db 214 GSSSGNTASLTITGAQAED-----ADYYCNSRDSGNHVVFGGQTKLTVLGAAGAEQKLISE 270  
QY 402 DDGEGR 408  
Db 271 EDLNGAA 277

RESULT 14  
US-10-374-780A-1400  
; Sequence 1400, Application US/10374780A  
; Publication No. US20040019927A1  
; GENERAL INFORMATION:  
; APPLICANT: Sherman, Bradley K  
; APPLICANT: Riechmann, Jose Luis  
; APPLICANT: Jiang, Cai-Zhong  
; APPLICANT: Heard, Jacqueline E  
; APPLICANT: Haake, Volker  
; APPLICANT: Creelman, Robert A  
; APPLICANT: Ratcliffe, Oliver  
; APPLICANT: Adam, Luc J  
; APPLICANT: Reuber, T. Lynne  
; APPLICANT: Keddle, James E  
; APPLICANT: Broun, Pierre E  
; APPLICANT: Pilgrim, Marsha L  
; APPLICANT: Dubell III, Arnold T

APPLICANT: Pineda, Omaira  
APPLICANT: YU, Guo-Liang  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS  
FILE REFERENCE: MBI-0047 CIP  
CURRENT APPLICATION NUMBER: US/10/374,780A  
CURRENT FILING DATE: 2003-02-25  
PRIOR APPLICATION NUMBER: 09/837,944  
PRIOR FILING DATE: 2001-04-18  
PRIOR APPLICATION NUMBER: 60/310,847  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 09/934,455  
PRIOR FILING DATE: 2001-08-22  
PRIOR APPLICATION NUMBER: 60/336,049  
PRIOR FILING DATE: 2001-11-19  
PRIOR APPLICATION NUMBER: 60/338,692  
PRIOR FILING DATE: 2001-12-11  
PRIOR APPLICATION NUMBER: 10/171,468  
PRIOR FILING DATE: 2002-06-14  
PRIOR APPLICATION NUMBER: 10/225,066  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,067  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: 10/225,068  
PRIOR FILING DATE: 2002-08-09  
NUMBER OF SEQ ID NOS: 2906  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 1400  
LENGTH: 323  
TYPE: PRT  
ORGANISM: Oryza sativa  
FEATURE:  
OTHER INFORMATION: Orthologous to G1073  
US-10-374-780A-1400

Query Match 4.2%; Score 122; DB 15; Length 323;  
Best Local Similarity 23.0%; Pred. No. 0.018;  
Matches 87; Conservative 37; Mismatches 122; Indels 132; Gaps 18;  
QY 200 PDGVAGAVSSRGDAGRVTVRKTSRSEAAASHSHSQTMOPR-----VGNLSGVEI 255  
DB 30 PDDTAMAG-MDPGGGGAG-----AGSR-----YFHLLRQOPSPPLSPLPTSHVKM 76  
QY 256 YLSQSSRNPTPRGSSFNHADPFNTVGAAGKGGGAAGDE-----KGACGG-G 302  
DB 77 EHSKMSDPKSPVREG-DH-----AGSGSGGVGGDHPSSAMVPEVGGSGSAGSG 127  
QY 303 GGHSPQPOAVAPAKRKDLHMLVWSSASPVSEAAVHVGAGADHADVLAKGAYD 362  
DB 128 SGGPTRRPRRPPGSKNPKPPIITVTRDSPNALHSHVLEV--AGADVDCVA----- 178  
QY 363 EYGRDDYSSRTKNG-----SGGADKGGPTLSKLSNSTAQLYPKDDGEGRAAAVAMPASV 418  
DB 179 -----EYARRRGRGVCLSGGAVVVALRQPG-----ASPPGSM 213  
QY 419 MTRILIMVWRKLRNPNNTYSLLGVVWSLVSYRWGTEMPAIIARSISILSDAGLGMAMF 478  
DB 214 VATL-----RGRFEILSLGTG-----LPPAPPAGAS-----GLTVF 245  
QY 479 SLGLFMALQPRIACGNKLAIAIMGVRFVAGPAPVMAAASIA-----VGLRGVLLHIA-- 530  
DB 246 LSG-----GGQVIGSVVGPLVAGPVLVMAASFANAYVRLPLEGEEVEVAP 295  
QY 531 -----IVQAALPQG 539  
DB 296 AAGGEAQDVAQAGPQG 313

RESULT 15  
US-09-738-626-6032  
Sequence 6032, Application US/09738626  
Publication No. US20020197605A1  
GENERAL INFORMATION:  
APPLICANT: NAKAGAWA, SATOSHI

APPLICANT: MIZOGUCHI, HIROSHI  
APPLICANT: ANDO, SEIKO  
APPLICANT: HAYASHI, MIKIRO  
APPLICANT: OCHIAI, KEIKO  
APPLICANT: YOKOI, HARUHIKO  
APPLICANT: TATEISHI, NAKO  
APPLICANT: SENOH, AKIHIRO  
APPLICANT: IKEDA, MASATO  
APPLICANT: OZAKI, AKIO  
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
FILE REFERENCE: 249-125  
CURRENT APPLICATION NUMBER: US/09/738,626  
CURRENT FILING DATE: 2000-12-18  
PRIOR APPLICATION NUMBER: JP 99/377484  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: JP 00/159162  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: JP 00/280988  
PRIOR FILING DATE: 2000-08-03  
NUMBER OF SEQ ID NOS: 7059  
SOFTWARE: PatentIn ver. 3.0  
SEQ ID NO 6032  
LENGTH: 585  
TYPE: PRT  
ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6032

Query Match 4.2%; Score 121.5; DB 9; Length 585;  
Best Local Similarity 22.4%; Pred. No. 0.046;  
Matches 82; Conservative 50; Mismatches 147; Indels 87; Gaps 15;  
QY 65 FAMNLRFLAAD--TLQKVAV-----LALLALASRLSSP-RALGLDWSITLFSLTPLNT 116  
DB 209 FAANAKMGASVGTNLRIYMGHERMQVMAATSIIKALDPVQRLAAERAFASFQATFOAD 268  
QY 117 LVWGIP---LLRGMGAS-SAGTLMVQVVVLCIIWYILMLFLFYRAARALVLDQF--- 169  
DB 269 VLTGMPPVSNLMQMGANGSAGEI-----ALGMDIEIAGS 302  
QY 170 -----PDGAAASIVSFRVDSVVSRLARGDVLEBAEPDGVAGAGAVSSRGDAGR 219  
DB 303 GQANSAAGLTPSGAAQGGVA-----NAGSIAPDAAVQGAAGQSGVSGFTVTDQL 352  
QY 220 -RTVVKSTSRSEAAASHSHSQTMOPRVSNLSGVEIYLSQSSRNPTPRGSSFNHADPFN 278  
DB 353 DGINIGDMLTSAASAGQSLANGLAMFTSTPNSASGAIPSSMSAASPL-----GAFGS 404  
QY 279 IVGAAGKGG-----GAAGDEEKACGGGGGHSFPQOAVAVPAKRKDLHMLVWSSAS 332  
DB 405 GAGLGAQGGSIGSSAPGALISSRAAGSAGSVPMGTGGPGA---PGITSDSLMGARTHGAS 461  
QY 333 PVSEAAVHVFAGGAGHADVLAKGAYDEYGRDDYS---SRTKNGSG---GADKGGP 385  
DB 462 SAGAVAPM-MGGAGGMSGGVVGAGGTGSGSKYARQTGSSVSSSSQSGSLGMVSGSGKP 520  
QY 386 TLSKLG 391  
DB 521 SISNFG 526

Search completed: March 3, 2004, 09:14:08  
Job time : 56 secs